Multi-Dataset Image Reconstruction for Longitudinal and Multi-Tracer Positron Emission Tomography

Ellis, Sam

Awarding institution:
King's College London

Multi-Dataset Image Reconstruction for Longitudinal and Multi-Tracer Positron Emission Tomography

Sam Ellis

Supervised by:
Professor Andrew J Reader
Professor Julia Schnabel

Thesis submitted to King's College London in fulfilment
of the degree Doctor of Philosophy

September 2018

Abstract

Repeated scanning of subjects is commonplace in clinical and research positron emission tomography (PET), either for the assessment of physiological change, or to provide complementary information from distinct biological processes. In many of these cases, the intrinsically noisy nature of PET data can hinder the interpretation of images, or necessitate the use of higher radioactive doses. Despite the existence of common information between repeated PET scans, they are typically reconstructed individually, representing a sub-optimal use of the available data. This thesis proposes and evaluates methods of transferring information between these multi-dataset PET images during the image reconstruction process, in order to improve the quality of the resulting images. Two approaches for sharing information are investigated, simultaneous reconstruction for longitudinal imaging and guided reconstruction for multi-tracer imaging.

For the simultaneous case, longitudinal datasets are coupled in the image reconstruction optimisation problem by using difference-image priors, thereby encouraging desired characteristics of the resulting difference images. This approach was applied to longitudinal PET imaging for tumour monitoring, showing reduced image noise and lower reconstruction errors compared to standard image reconstruction methods. Importantly, these methods denoise in the difference image domain, thereby providing longitudinal images that exhibit noise levels comparable to standard PET images reconstructed with higher levels of counts. In general, using the proposed methods to reconstruct S scans results in noise levels similar to those usually achieved by using traditional reconstruction methods with datasets containing S times the number of counts. Experimental results demonstrated how changing the characteristics encouraged in difference images can affect the reconstruction of tumour regions in the resulting PET images, highlighting the need for application-specific selection and validation of difference-image priors.

For multi-tracer PET imaging, guided reconstruction was proposed as a

method of transferring information from high quality PET images to lower quality images. Inspired by the literature on anatomically guided PET image reconstruction, a weighted quadratic penalty was applied to [^{18}F]-fluorodeoxyglucose (FDG)/[^{11}C]-methionine imaging of brain tumours. Image structure information was extracted from the [^{18}F]-FDG images and embedded into the reconstruction of the [^{11}C]-methionine images. Guiding [^{11}C]-methionine image reconstruction with [^{18}F]-FDG information in this way led to edge-preserving noise reduction in the [^{11}C]-methionine images compared to standard reconstruction methods. Furthermore, guiding reconstructions using [^{18}F]-FDG images was observed to outperform anatomically guided PET image reconstruction in some regions. These results suggest that multi-tracer PET imaging could benefit from the use of guided image reconstruction methods, with many possible applications for the utilisation of the guided approach.

Overall, this thesis demonstrates the principle of sharing information between PET images during reconstruction, showing improved image quality when applied to specific contexts. Future work remains in further assessment of the proposed methods in clinical contexts, and testing of the methods in other related areas of multi-dataset PET imaging.

Acknowledgements

Firstly, I would like to thank my supervisor Andrew Reader for his continuous help and useful advice throughout my PhD studies. I would also like to thank Julia Schnabel, Paul Aljabar, Andy King, Paul Marsden, and Lefteris Livieratos for their productive comments and support.

I am grateful to my colleagues at King's who have collaborated with me and who have helped me to achieve this thesis. In particular, thanks to Martín Belzunce and Abolfazl Mehranian for their invaluable insights (and code) for using clinical PET data, and to Gary Cook, Colm McGinnity, Andrew Mallia, and Zach Chalampalakis for providing me with such data.

As well as colleagues, I have made some great friends during my PhD. Thank you to Chris, Patrick, and Dan, with whom I have been able to talk about research, but also about life in general.

Thanks also to my friends outside of King's over these years, Rob, Chris, Becky, and Jenni, for their support and for the good times that have kept me connected to the real world beyond the PhD.

I also owe a great deal to Camila, without whose support I'm not sure I would've finished this PhD, and who is always challenging me to improve, both professionally and personally. Thank you very much.

Last, but far from least, I would like to thank my family, Mum, Dad, Jamie, Mark, and everyone else, for getting me to the point of studying for a PhD in the first place. I'm sure that it would not have been possible without the encouraging environment that I grew up in.

This PhD position was funded by the King's College London and Imperial College London EPSRC Centre for Doctoral Training in Medical Imaging (grant number EP/L015226/1).

Contents

List of Figures

List of Tables

Chapter 1: Introduction to PET Imaging and the Uses of Multi-Dataset PET

1.1 Introduction to PET Imaging

Positron emission tomography (PET) is an emission tomography (ET) medical imaging modality that is used to map the distribution of radioactively labelled molecules *in vivo*. By incorporating positron emitting radionuclides into biologically active molecules, various biochemical and physiological processes can be targeted for observation and quantification. In general, PET is highly sensitive to small quantities of radioactivity, able to detect activity concentrations within the body down to the picomolar scale [1]. This high sensitivity allows trace quantities of radioactivity to be observed, avoiding perturbation of the biological systems under consideration.

When the radioactively labelled tracer (or radiotracer) is introduced into the body, most commonly by intravenous injection, the anatomical region to be imaged is positioned within the scanner, which for PET typically consists of a set of static rings of radiation detector elements. Over the scan duration, numerous radioactive events are recorded in these detector systems, which allow a map, or image, of radioactivity concentration to be produced by a process known as image reconstruction. These images usually have units of

activity per unit volume ($\mathrm{Bq{\cdot}ml^{-1}}$). From these values, the semi-quantitative measure known as the standardised uptake value (SUV) can be calculated, as given by [2]:

$$\mathrm{SUV} = \frac{a}{(d/w)} \tag{1.1}$$

where a is the observed activity concentration, d is the decay-corrected injected activity and w is the weight of the subject being scanned. The unit of SUV is $\mathrm{g{\cdot}ml^{-1}}$. SUVs allow for improved inter-scan and inter-subject comparison of biological activity compared to radioactivity concentration, which depends, among other things, on injected activity levels and the size of the subject.

The most common clinical application of PET imaging is in oncology. Tumours often have considerably different functional and/or metabolic characteristics compared to healthy tissue, making direct observation of these processes via PET imaging a valuable tool. In particular, tumours typically exhibit high levels of glucose metabolism, which they require to proliferate and maintain growth [3]. For this reason, the most commonly used PET radiotracer for oncology purposes is fluorine-18 (^{18}F) labelled fluorodeoxyglucose (FDG) [4]. FDG is imported into cells from the blood supply by glucose transporters. Once within cells, FDG is phosphorylated by the hexokinase enzyme. At this stage, normal glucose is then further metabolised to produce energy for the cell, but FDG becomes trapped, remaining in the cell for a prolonged period of time. By labelling FDG with ^{18}F and performing imaging after a sufficient time interval, the majority of the remaining [^{18}F]-FDG is expected to be distributed according to glucose metabolism throughout the body. Tissues that use high levels of glucose (for example cancer cells) exhibit a higher activity concentration, corresponding to high intensity regions in the reconstructed PET images.

Despite its wide usage in oncology, [^{18}F]-FDG has the disadvantage of being non-tumour-specific. Inflammation, fat, muscle, and cortical tissue can all exhibit high [^{18}F]-FDG uptake, potentially leading to false positives in cancer detection [5–10]. Furthermore, some cancerous tissues exhibit normal levels of glucose metabolism, and so could be missed using [^{18}F]-FDG imaging [6, 11].

For these reasons, alternative oncology PET radiotracers have been developed, targeting a wide range of biological processes [12–15]. Among the most widely used of these alternative PET radiotracers are those based on protein synthesis and transport, for example carbon-11 labelled methionine ([^{11}C]-methionine) [16–18], ^{18}F-labelled thymidine [19, 20], or ^{18}F-labelled tyrosine [21–23]. Other popular alternative PET radiotracers include ^{18}F-labelled sodium fluoride for detection of bone metastases [24] and ^{11}C- or ^{18}F- labelled choline for the detection and staging of prostate cancer [25].

PET has also been widely used for brain imaging. Due to the high glucose metabolism of healthy cerebral grey matter, [^{18}F]-FDG PET has been used to explore cerebral metabolic activity in neurodegenerative diseases such as Alzheimer's disease and dementia [26–28]. In addition, since β-amyloid accumulation in the brain is known to be an indicator of Alzheimer's disease [29, 30], there has been recent development of β-amyloid targeting PET radiotracers such as the ^{11}C-labelled Pittsburgh compound B (PiB) [31, 32], [^{18}F]-florbetapir [33] and [^{18}F]-florbetaben [34]. This is a particularly active area of research given the growing prevalence and impact of such diseases in the developed world [35].

Another established use of PET in brain imaging is in the study of receptor-ligand systems within the brain, for example in the dopamine receptor system [36–38]. PET is a popular tool for such studies due to a) the previously mentioned high sensitivity of PET, which allows observation of the system without perturbation; b) the large range of PET radiotracers available for targeting different receptor-ligand systems [39]; and c) the quantitative nature of PET, which allows for *in vivo* calculation of parameters of interest such as binding potentials, rate constants, and volumes of distribution [40].

Finally, there has been an increased interest in using PET for cardiovascular imaging, where the most common objective is to assess myocardial perfusion and/or myocardial viability, both for non-invasive diagnosis of coronary artery disease and prior to performing re-vascularisation therapy following infarction or ischemia [41, 42]. [^{18}F]-FDG is widely used as a cardiac PET radiotracer for myocardial viability, due to the high glucose metabolism of myocardial tissue.

Rubidium-82 (^{82}Rb), oxygen-15 (^{15}O) labelled water, and nitrogen-13 (^{13}N) labelled ammonia are the most common perfusion tracers for cardiac PET [42, 43]. Novel radiotracers that target specific mechanism of cardiovascular disease have recently been developed. For instance, ^{18}F-labelled sodium fluoride has been used for the imaging of coronary artery plaques, targeting the growth of micro-calcifications that occur within atherosclerotic plaques [44–46].

1.2 Multi-Dataset PET

In many applications there is a need to collate information from multiple PET scans to obtain information that it is not possible to observe with a single scan. This repeated observation using PET is referred to as multi-dataset PET in this thesis. There are two broad categories of multi-dataset PET: the observation of changes in one aspect of physiology using repeated PET scans with a single radiotracer, and the use of more than one radiotracer to obtain complementary information from different physiological processes, known as multi-tracer PET.

1.2.1 Detection and Quantification of Change in PET

Observing or measuring the variation of a specific biological process throughout time is an important use of PET imaging. These contexts can further be categorised as either 'longitudinal', where scans are separated by a period of time in which the biological system is left to evolve naturally, and 'sequential', in which scans are repeated in quick succession following an artificially introduced stimulus designed to affect the physiology under examination. The following is a brief review of areas where longitudinal and sequential PET examinations are used, with a particular focus the longitudinal assessment of treatment response in oncology.

Longitudinal PET imaging is an important tool for the study of neurodegenerative diseases. Throughout the onset of diseases such as Alzheimer's disease there are changes in the functional behaviour of the brain. For exam-

ple, it has been demonstrated that there is a reduction in grey matter glucose metabolism that correlates with worsening Alzheimer's disease [27]. In addition, as mentioned in Section 1.1, there are other radiotracers that have been designed to image more specific markers of neurodegeneration, such as [^{11}C]-PiB for amyloid levels, which are known to change early in the onset of Alzheimer's disease [30]. These amyloid levels can be imaged longitudinally [47] to observe neurodegeneration, and have been shown to correlate with the reduction in glucose metabolism [48]. Furthermore, measurement of amyloid levels by PET imaging after treatment may provide early feedback about the efficacy of that treatment [49].

PET can also be used to measure short-term changes in brain activity induced deliberately, for example using visual stimuli [50], behavioural tasks [51], pain administration [52], music [53], or psychoactive drugs [54, 55]. One of the common targets for such studies is the dopaminergic pathway, which is important in many brain disorders [51, 52, 56, 57]. Quantitative measurement of the effect of psychoactive drugs on brain function in terms of biological parameters of interest allows an improved understanding of the underlying mechanisms of such disorders, and can inform the design of future treatments. Note that although these methods are often performed in a single scanning session they can be considered as multi-dataset PET insofar as they comprise multiple observations/data acquisitions that are independently processed and then compared. In these cases, statistical parametric mapping methods are often employed to perform analysis of the changes observed. These methods use rigorous statistical methods and voxel-level hypothesis testing to determine which areas of the brain have changed significantly between sequential scans based on multiple observations (subjects and/or experiment repetitions) [58]. This allows the localisation and quantification of statistically significant changes to be performed.

Another application of multi-dataset PET is the assessment of myocardial perfusion under rest and stress conditions, using perfusion radiotracers such ^{82}Rb, [^{15}O]-water, or [^{13}N]-ammonia. In coronary artery disease, myocardial blood flow can be reduced regionally under stress conditions, with the magni-

tude of stress-induced defect related to the severity of pathology [43]. Stress can be induced by the use of exercise regimes [59, 60] or, more commonly, pharmacologically via the administration of adenosine [61]. In these protocols, patients can be scanned in any order (i.e. stress followed by rest [43], or vice versa [42]), with a time interval between scans of ∼10 min. Similarly to the cases mentioned above, these rest and stress datasets are usually reconstructed independently and then compared.

Nonetheless, one of the most important examples of the use of longitudinal PET is in cancer imaging, where the response or progression of a tumour during or after treatment is often of high interest [62]. Whilst in clinical contexts the recommended measure of the response or progression of a tumour to therapy has been tumour size as measured using x-ray computed tomography (CT) [63], PET has proven to be a valuable alternative due to the fact that functional changes often precede anatomical ones [64]. This potential earlier assessment of treatment efficacy allows quicker identification of responders and non-responders [65–67], facilitating enhanced patient management by adapting treatment where appropriate [68]. Observation of longitudinal changes in tumours is also useful for assessing the effectiveness of new biomarkers in reflecting tumour response [69]. Longitudinal PET studies have been performed for various cancer treatments, such as chemotherapy [70–72], radiotherapy [73], and immunotherapy [74]. In studies where PET has been used for assessing oncology treatment response, the number of scans that are performed varies greatly, ranging typically from two to five scans [66, 67, 73], up to more than ten in some specific studies [74].

In these studies, assessment of change is usually quantified by the change in SUV measured within a tumour, reflecting the clinical standard for reporting on individual PET images. There are however a number of ways to calculate tumour SUVs, differentiated by the set of voxels that are aggregated and the method used to calculate a single representative SUV [2]. Studies have used many methods, including the mean SUV in a region-of-interest (ROI) with a pre-defined shape (e.g. a spherical) [65, 71, 75–77], the mean in an isocontour-defined ROI [78], the mean in a manually delineated tumour ROI [72], and

the maximum recorded SUV within the tumour [70, 79, 80]. Another method that has been proposed is the use of SUV peak values calculated by finding the mean SUV in a standardised ROI centred at the maximum SUV voxel [64]. The use of differing ROIs for SUV measurement has been shown to dramatically affect quantification in treatment response PET scans [81], a fact attributed to tumour heterogeneity and varying levels of robustness to variations in image quality.

Less commonly, tumour response may be quantified with PET in terms of the change of the underlying biochemical parameters such as enzymatic kinetic rates [66, 72, 82–85]. Using such measures in place of semi-quantitative SUVs is more representative of the underlying biological processes than SUV measures. However, these methods require prolonged scan times to sample the radioactivity distribution through time, generally precluding them from clinical settings where patient comfort and throughput are key considerations.

Recent research has proposed more advanced methods for assessing the longitudinal change in tumour activity based on PET scans, using advanced image processing techniques. Such methods include the classification of individual voxels into distinct groups of change (e.g. response, stable, progression) [86–90] and advanced metrics of tumour change calculated using all relevant voxels directly rather than using aggregated values like SUV [91]. Figure 1.1 shows an example of one such method [86], where green voxels in the parametric response image (Figure 1.1(c)) indicate voxels determined to have responded postively to treatment. However, these techniques require accurate alignment of the follow-up and baseline images, particularly in methods where change is calculated on the voxel-level. For this reason methodologies for evaluating the suitability of image registration methods for voxel-level analysis of longitudinal PET changes are important [92].

1.2.2 Multi-Tracer PET

In addition to multi-dataset PET for the repeat observation of specific biochemical processes, PET can also be repeated to obtain complementary infor-

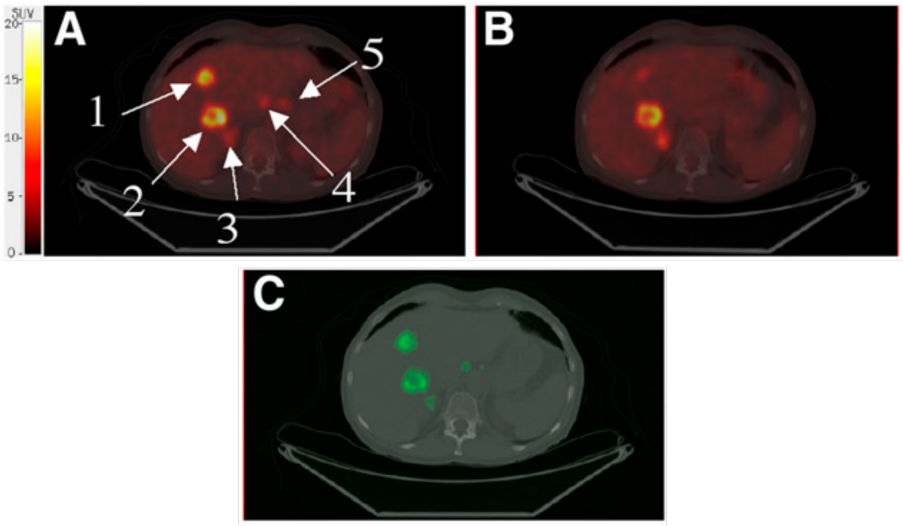


Figure 1.1: Example of tumour response as observed with PET imaging using voxel-wise SUV changes, overlaid on anatomical CT images. The baseline tumour activity, indicated with arrows in (a), reduced in the follow-up examination (b). The regions of change can be visualised separately (c). This figure is an adaptation of a figure originally published in Necib *et al.* 'Detection and Characterization of Tumor Changes in ^{18}F-FDG PET Patient Monitoring Using Parametric Imaging'. J Nucl Med. 2011;52:354-361 [86] © SNMMI.

mation about different processes by using multiple radiotracers (multi-tracer PET). The simplest way to perform multi-tracer PET is to perform separate scans for each radiotracer. In this case the scans can be performed as normal, with the only necessary consideration being the possibility of contamination of acquired data with signal arising from previously administered radiotracers, an issue that can be addressed by using shorter lived radiotracers in earlier scans, allowing shorter time intervals between the scans. In contrast, simultaneous multi-tracer PET is more challenging due to the difficulty in separating the signals from different radiotracers, a topic of ongoing research [93–97].

Multi-tracer PET is of interest in oncology, where different cancer phenotypes may be discernible based on their characteristics as observed using multiple PET radiotracers. For example, in the imaging of brain tumours, the complementary information obtained from PET images using [^{18}F]-FDG for glucose metabolism imaging and a protein synthesis/transport radiotracer such as [^{11}C]-methionine can improve the detection and grading of malignant gliomas [17, 98]. Other suggested applications in oncology include the use of

complementary information from [^{18}F]-FDG, [^{11}C]-acetate, and [^{18}F]- or [^{11}C]-labelled choline for evaluation of prostate cancer [99], and [^{18}F]-FDG and [^{11}C]-acetate for liver cancer [100].

Multi-tracer PET is also of potential importance in brain imaging due to the interconnected nature of biochemical pathways within the brain. For example, as mentioned above, both glucose metabolism and amyloid load are of interest in the monitoring of neurodegeneration [48]. In brain activation studies, there is sometimes interest in observing brain perfusion using [^{15}O]-water and glucose metabolism with [^{18}F]-FDG [95]. Another example is the observation of glucose metabolism and dopaminergic pathways via the use of [^{18}F]-FDG and [^{11}C]-raclopride in the differential diagnosis of Parkinson's disease, progressive supranuclear palsy, and multiple system atrophy [96]. One final example of multi-tracer brain PET is the observation of the the glutamate and dopamine receptor systems, an area of importance in various neurological conditions [101, 102].

1.3 Motivation and Plan of this Thesis

As outlined above, there are many contexts in which subjects are scanned multiple times using PET imaging, whether to observe longitudinal variations of pathological physiology, changes in physiology due to deliberately induced stimuli, or complementary physiological processes using different PET radiotracers. In all of these cases, since the scan subject remains the same, some level of the information contained within the scans is redundant. Nonetheless, the conventional approach for processing and reconstructing images from these datasets is to reconstruct each image separately, neglecting to account for the common information contained within the images. Therefore, there is an opportunity to improve such multi-dataset PET images by transferring information between scans during the image reconstruction process. Such a transfer of information between image reconstruction problems has been explored before in the past in other medical imaging contexts (see Section 2.4), but to the best of the author's knowledge such methods have not previously

been applied to multi-dataset PET.

This thesis proposes image reconstruction methods that take advantage of this similarity between longitudinal and multi-tracer PET images, and evaluates the performance of these methods using simulation studies and real data experiments. There are two main approaches to the multi-dataset PET image reconstruction problem that are proposed in this thesis: simultaneous reconstruction of longitudinal oncology treatment response scans, and guided image reconstruction of multi-tracer PET images. In the simultaneous reconstruction case, longitudinal images are reconstructed jointly, allowing information to be shared between all images during the reconstruction. In the multi-tracer guided reconstruction case, one radiotracer image is reconstructed, and information extracted from that image is introduced into the reconstruction of another radiotracer image. The organisation of this thesis is as follows.

Chapter 2 describes the physical processes and technical considerations of PET data acquisition, and provides the theory behind statistical PET image reconstruction, including methods for incorporating prior expectations into the image reconstruction problem. Related work in multi-dataset image reconstruction in other contexts is then reviewed to motivate the methods proposed in this thesis.

Chapter 3 proposes a method of performing the simultaneous reconstruction of pairs of longitudinal oncology treatment response PET images that takes advantage of the expected sparse nature of the resulting difference images. By using penalised PET image reconstruction methods that encourage sparsity in the difference image domain, PET data is effectively shared between the two scans, allowing for improved image quality. Two forms of the sparsity-encouraging penalty are presented, and the performance of these methods was evaluated using 2D simulation studies and experiments using inserted lesions in 3D real datasets.

Chapter 4 generalises the methods proposed in Chapter 3, by expanding the image reconstruction problem to include an arbitrary number of longitudinal scans, including a factor to correct for the varying activity levels across longitudinal images, and accounting for misalignment between scans. Further-

more, two alternatives to the sparsity-encouraging penalty are proposed that a) encourage difference images with low entropy, and b) encourage the formation of difference images with sparse spatial gradients. 2D simulation studies were performed to assess the performance of the competing penalty terms in both two- and five- scan cases. A two-scan clinical data case study was carried out to explore the performance of the methods in a real data situation.

Chapter 5 focuses on guided multi-tracer image reconstruction, exploring the use of [^{18}F]-FDG images to guide the reconstruction of [^{11}C]-methionine images in brain oncology. The *a priori* similarity between image voxels is calculated based on the [^{18}F]-FDG image using patch-based measures, and this information is included in the [^{11}C]-methionine image reconstruction process. Two competing methods for including this information are used, consisting of a penalised reconstruction approach and a post-reconstruction guided filter. The results of 3D simulation studies that explore the effect of a series of hyperparameters on the quality of reconstructed images are presented. The proposed method is then applied to a real [^{18}F]-FDG/[^{11}C]-methionine data case.

Chapters 6 and 7 then discuss the contributions of this thesis in the context of the wider literature on PET image reconstruction and multi-dataset image reconstruction methods, suggest possible areas for future work, and make concluding remarks on the work presented in this thesis.

Chapter 2:
PET Data Acquisition and Image Reconstruction

This Chapter describes the physical processes involved in PET data acquisition, an overview of the processing of PET data, and the formulation and solution of the statistical PET image reconstruction problem. Related work in the area of multi-dataset image reconstruction is then reviewed. This Chapter provides the framework in which the image reconstruction methods proposed in this thesis are defined.

2.1 PET Data Acquisition

Inside the body, a radionuclide attached to a PET tracer molecule will undergo radioactive decay by emitting a positron (as well as an undetectable neutrino). This positron travels a small distance inside the surrounding tissue (Figure 2.1(a)), with the expected distance dependent upon the original energy of the positron and the properties of the tissue. For example, positrons produced by ^{18}F have a root mean square range in water of 0.23 mm, and for ^{11}C the corresponding value is 0.39 mm [103, 104]. At the end of its path, the positron interacts with an electron in the tissue by annihilation, producing two back-to-back 511 keV photons according to Einstein's well-known $E = mc^2$ (Figure 2.1(b)). By placing the patient inside the ring (or multiple rings) of radiation detectors that form a PET scanner, these two high-energy photons can be observed (Figure 2.1(c)), tracing a line-of-response (LOR) that contains the

position of the decay event. This LOR can be characterised by its distance r and angle θ relative to a Cartesian axis. The observed event is recorded by adding a count to the appropriate bin in a discretised r, θ space (Figure 2.1(d)). By observing and recording millions of such counts, a PET dataset is built up over a time period that typically ranges from a few minutes up to an hour. Such PET datasets are known as 'sinograms', since placing a point source emitter at an arbitrary position inside a PET scanner results in a sinusoidal trace in the data. The principle is readily extendible to 3D PET imaging, by concatenating detector rings to form a cylindrical bore, and to 4D PET by performing additional temporal binning.

Due to the finite speed of light, when the original radioactive decay site is not at the midpoint of the LOR there is a non-zero time interval between the arrival of each of the 511 keV photons. This delay depends on the diameter of the PET detector ring, and for conventional PET scanners is on the nanosecond scale. Therefore, in order to detect coincidences, a time interval is required so that if a second photon is detected within the interval τ from an initial detection, the two are considered to have resulted from the same radioactive decay event. The LOR connecting the two detections is then calculated, and the event is binned into sinogram space as mentioned above.

In fact, if detectors of a suitable timing resolution are used, the observed events can also be binned according to the delay between the arrivals of each photon, restricting the segment of the LOR along which the original decay would be expected to have occurred. The use of such delay-binned data is known as time-of-flight PET [105], and although the remainder of this work considers only non-time-of-flight PET, many of the ideas and methods presented here are extendible to the time-of-flight case.

As described above, the simplest form of non-time-of-flight PET data is the sinogram (or a set of sinograms in 3D acquisitions), which records just the number of counts recorded in each r, θ bin throughout the acquisition. Nonetheless, other PET data formats are available, representing different levels of processing of the data. The most raw form of PET data consists of time-stamped single photon detections within the individual detector elements, along with

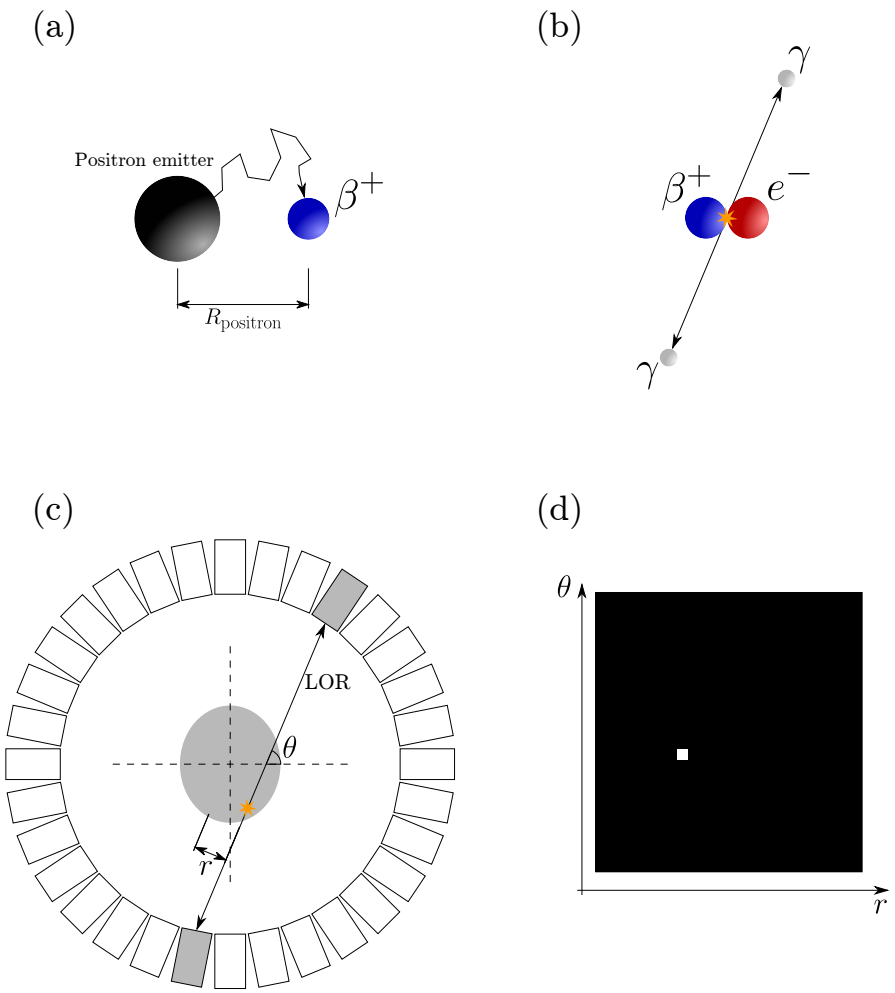


Figure 2.1: Principle of PET imaging. (a) The tracer radionuclide emits a positron (β^+ particle), which travels an effective distance of R_{positron} from the decay site along a random path, determined by its interactions with other matter. (b) At the end of this path, the positron meets an electron (e^-) and the two particles annihilate, producing two back-to-back 511 keV photons (γ). (c) When the patient is placed in scanner comprising a ring of radiation detectors, these photons can be observed from signals recorded in a pair of detectors (shaded blocks). This coincidence detection is characterised by its line of response (LOR), parameterised by its observed angle (θ) and distance (r) relative to a Cartesian axis (dashed lines). (d) The detected coincidence is then recorded by adding a count to the corresponding location in the discretised 'sinogram' space. A typical PET acquisition involves the observation of millions of counts per 2D slice. Note that while this description applies specifically to 2D static PET, an analogous process applies to 3D and dynamic PET.

spatial information about the photon-detector interaction if this is available. The next level is list-mode data, which contains a time-stamped list of all recorded coincidences, i.e. a list of all *pairs* of single photon detections that occurred within the coincidence timing window τ for each detector pair combination. Above this is sinogram PET data, where the coincidences contained in list-mode data are aggregated over a time period, resulting in a simple list of number of observed coincidences for every detector pair. Dynamic PET, which aims to achieve a temporal resolution of PET data, can use either list-mode data or a time-series of sinograms produced by binning list-mode data into a set of time-frames of arbitrary duration.

2.1.1 Degrading Effects in PET Imaging

There are a number of physical processes that occur during PET data acquisition that degrade the data via incorrect observation of counts or by introducing uncertainties into the locations of decay events. These processes result in degradation of PET image quality if not compensated for, affecting image contrast, noise levels and quantitative accuracy. The following description briefly reviews these effects, but a more complete discussion can be found in [104] and in references below.

Photon Attenuation As with any photons travelling through matter, some of the high-energy annihilation photons in PET are attenuated, for example by the tissue of the patient being scanned. The probability of attenuation per unit length of photon path is known as the linear attenuation coefficient, and is dependent on the atomic number of the attenuating material and the energy of the photon in question. Typical linear attenuation coefficients encountered in PET imaging range from $\sim 0\,\mathrm{cm}^{-1}$ for air, through $0.096\,\mathrm{cm}^{-1}$ for water, and $0.172\,\mathrm{cm}^{-1}$ for bone [106]. Attenuated photons manifest in PET data as a lowering of recorded counts along a given LOR in accordance with the total linear attenuation coefficient of that LOR (Figure 2.2(a)). Since annihilation photons produced at the centre of the patient must traverse more

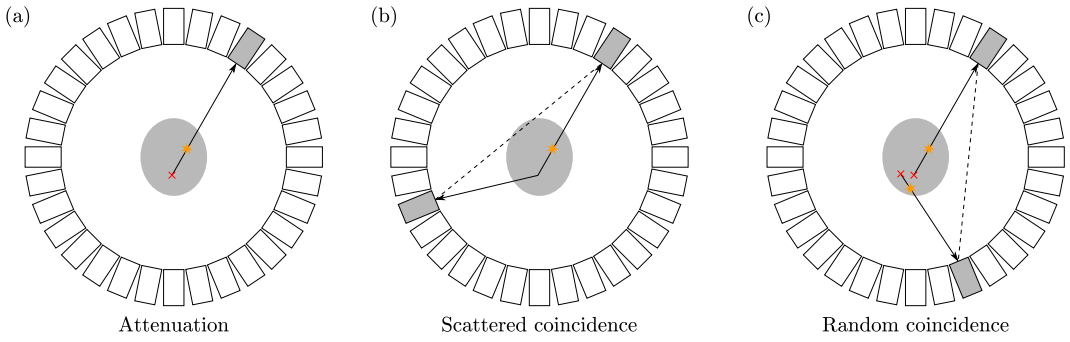


Figure 2.2: (a) Attenuated photons (red cross) prevent the detection of a radioactive decay event, lowering the rate of recorded counts in accordance with the attenuating properties of the patient/object in the field-of-view. (b) When photons undergo Compton scattering within the patient, an incorrect LOR can be assigned to the recorded coincidence (dashed line). (c) Detection of two photons from two separate positron emission events within the time window can result in incorrect LOR assignment (dashed line). Here is shown the case where one of each of the pairs of photons is attenuated, but random coincidences are also observed from unattenuated photons when the activity within the field-of-view is sufficiently high, or when one of a pair of photons does not interact with a detector.

highly-attenuating matter before leaving the body, they are attenuated proportionally more than photons produced near the edge of the patient, causing the characteristic appearance of non-attenuation-corrected PET images, with reduced counts from within the body and a 'hot-skin' effect [107]. To mitigate the effects of attenuation and to produce quantitative PET images, a map of attenuation coefficients throughout the field-of-view can be produced via a range of methods [106,108–114], and used to correct the PET images, either by correcting the data before reconstruction or by inclusion in the reconstruction process.

Scattered Coincidences One component of attenuation is scatter, whereby the direction of travel of a photon is changed and its energy reduced. In PET imaging this can cause the misplacement of a recorded LOR relative to the site of the original radioactive decay (Figure 2.2(b)). A proportion of scattered photon coincidences (also known as 'scatters') can be effectively ignored in the data acquisition process by using energy windowing methods, and there exist many methods of estimating the remaining counts contribution from scatters [115]. These detected scatters can be considered as an additive background

term contributing to the measured data (see Section 2.2.2), and so can be either pre-corrected prior to image reconstruction, or incorporated into the image reconstruction process.

Random Coincidences Due to the non-zero coincidence timing window τ required to record decay events that occur away from the centre of the field-of-view, it is possible that photons from two different decay events may be erroneously recorded as a single coincidence detection. Figure 2.2(c) shows the case where an accidental coincidence arises due to the attenuation of one photon from each of a near-simultaneous pair of decays, but in general these so-called 'random' coincidences (also known simply as 'randoms') can occur in any situation, with the probability of occurrence increasing quadratically with the activity level within the field-of-view. These random coincidences generally contribute an approximately uniform additive background to the measured sinogram data, since the directions of each of the photons are uncorrelated, thereby triggering each pair of detectors with equal probability (ignoring other effects like attenuation). As such, randoms can be corrected similarly to scatters, either via data pre-correction, or by inclusion in the image reconstruction algorithm.

Resolution Degradation and Limitations Spatial resolution in PET is limited by a number of factors of varying importance that depend on positron physics and detector limitations. The first is the movement of positrons within the tissue surrounding the decay event site prior to annihilation (Figure 2.3(a)) [116]. Since positrons are ejected from nuclei with a non-zero amount of energy, they can travel away from the emission site before annihilating with an electron. The expected value of this distance depends on both the nuclear structure of the emitting radionuclide in question, and on the properties of the surrounding tissue. The direction of positron movement is isotropic in a uniform medium under normal conditions (positrons in high magnetic fields such as those encountered in magnetic resonance imaging can be restricted in their movements), resulting in a corresponding isotropic point spread function (PSF)

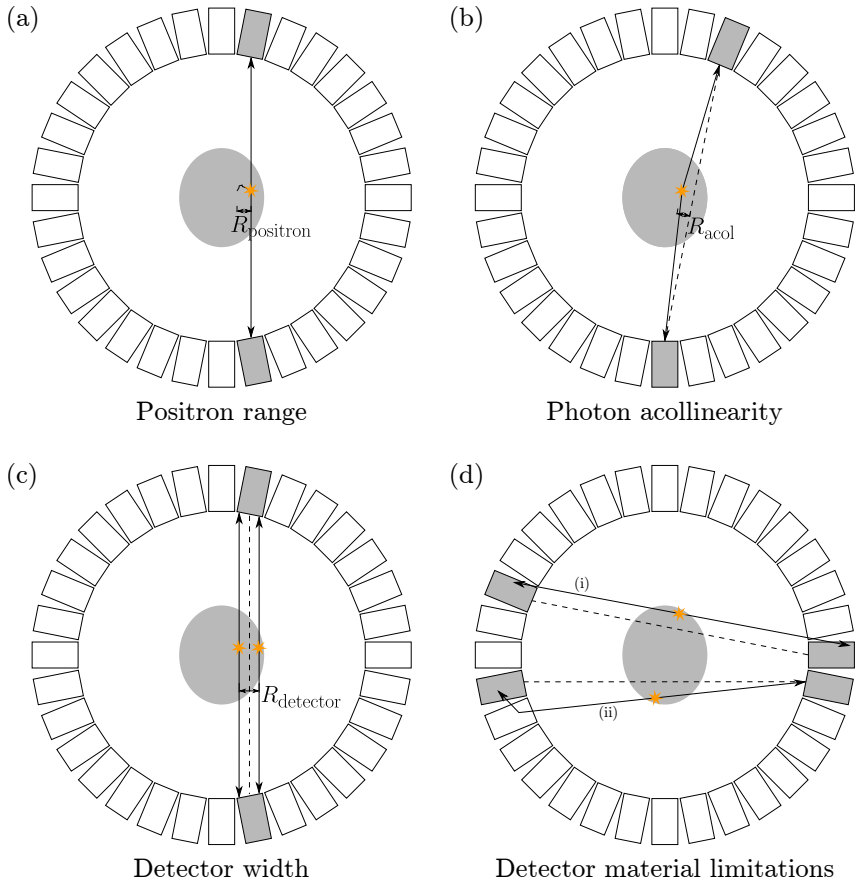


Figure 2.3: Causes of resolution degradation in PET imaging, often occurring due to misplaced recorded LORs (dashed lines). (a) Positron-electron annihilation rarely takes place in the exact location of the radioactive decay event due to the positron energy at the time of emission. The expected range of positron transit prior to annihilation (which is dependent on the radionuclide in question) causes a displacement in the recorded LOR for that recorded pair of photons. (b) At the point of annihilation, the positron-electron pair may have a non-zero momentum, causing acollinearity of the produced photons, displacing the observed LOR. (c) The finite width of detectors can lead to spatially distinct decay events being recorded as having occurred along the same LOR. (d) Other detector effects can further degrade resolution, for example: (i) depth-of-interaction effects where photons are not detected in the first detector along their path, or (ii) photon scattering within a detector element, resulting in an incorrect detector recording the photon.

due solely to positron range.

Another common resolution degradation effect is that of photon acollinearity. When the positron-electron pair possesses a non-zero total linear momentum prior to annihilation, the produced photons do not occur collinearly. Instead they get slightly deflected (with a range of about $\pm 0.25^\circ$ [104]) and travel in different directions, potentially causing incorrect LOR assignment as shown in Figure 2.3(b).

The third common cause of degraded PET image resolution is the physical width and spatial resolution of the detector elements (Figure 2.3(c)). Two decay events may be spatially separated and yet still be recorded as occurring along the same LOR. Note that while some detectors perform spatial localisation of the detection point within the detector element (e.g. [117]), there is always a minimum resolution of the position of LORs due to the uncertainty associated with this position estimation. In this way, all PET systems possess a spatial resolution component based on the detector geometry.

Beyond these fundamental resolution-degrading effects, detector materials can also further degrade resolution via insufficient stopping power allowing photons to pass through one detector element into another, or by allowing inter-detector scattering (Figure 2.3(d)). Both of these effects cause LORs to be misplaced, further degrading the intrinsic spatial resolution of the PET scanner.

PET image resolution is generally spatially variant as a result of all of these effects, with the highest resolution achieved in the centre of the field-of-view. Resolution effects can be corrected in a post-reconstruction manner, but more commonly are incorporated into the reconstruction algorithm [118]. Using resolution recovery methods such as these can produce sharper images, although quantification of such images can be affected by the presence of ringing artefacts [118].

Normalisation The overall sensitivity of a PET scanner refers to the proportion of true radioactive decays that it is able to detect, beyond the effects of attenuation. All PET scanners exhibit variations in sensitivity, both spatially

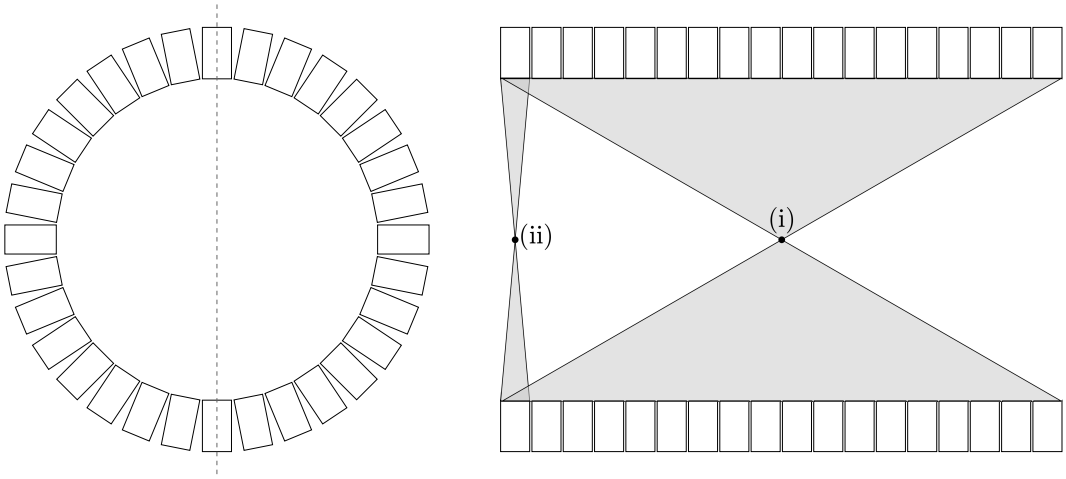


Figure 2.4: Considering the cross-section through a cylindrical PET scanner, it is apparent that the solid angle coverage (grey shaded area) is greater for a point source positioned in the centre of the axial field-of-view (i) compared to an identical point source placed close to the edge of the axial field-of-view (ii). This means that a greater proportion of emitted annihilation photon pairs are detected from the centre of the axial field-of-view, causing a spatially varying sensitivity that must be accounted for to produce quantitative PET images.

throughout the field-of-view and temporally [119, 120]. These variations arise from several distinct effects that are briefly described below. The process of correcting for these sensitivity variations is known as normalisation correction.

The geometric sensitivity of a scanner is the spatial variation in sensitivity due to the geometry of the gantry. In current clinical PET scanners this variation is strongest in the axial direction, where the solid angle coverage for a point source at the axial centre is much higher than for a point positioned at the edge of the axial field-of-view (Figure 2.4) [104, p. 333].

Another source of spatially varying sensitivity is variations in the efficiencies of the detectors that comprise the scanner [120]. Due to differing quantum efficiencies or photomultiplier gain levels, pairs of detectors will have unique values for the proportion of incident annihilation photons that are detected. This effect causes a spatial variation in sensitivity that is scanner-specific and time-dependent, requiring regular quality control scans to re-evaluate the performance of each detector element. In extreme cases, a detector element may fail completely, rendering a set of LORs as unobservable in the sinogram. A related sensitivity variation is due to the arrangement of detector elements;

it is common for there to be gaps between detector elements due to physical hardware limitations. In practice, these gaps act like failed detector elements, causing a spatially varying sensitivity that can be incorporated into the time-invariant geometric sensitivity.

Finally, sensitivity can vary temporally due to dead-time [121]. Dead-time refers to the time period after a detection during which a detector is unable to record further photons due to hardware limitations. There are two types of dead-time: paralysable and non-paralysable. Paralysable dead-time occurs when the arrival of a second photon within the dead-time resets the recovery time, for example by producing more scintillation photons in the crystal that must be cleared before the next annihilation photon can be observed. Non-paralysable dead-time, conversely, is a fixed amount of time between observable photons, independent of the arrival of any more during the dead-time period. Correcting for dead-time is difficult, since it cannot be known *a priori* when dead-time leads to unobserved coincidences. However, there are methods for estimating proportional losses due to dead-time and incorporating these as another normalisation component [120].

The total effect of these spatially and temporally varying sensitivity effects is to reduce the number of recorded coincidences non-uniformly throughout the field-of-view. Assuming that the proportion of unobserved counts can be estimated for each sinogram bin, normalisation sinograms which account for these losses can be produced.

2.2 Formulation of the Statistical PET Image Reconstruction Problem

Statistical PET image reconstruction refers to the production of PET images from acquired data using statistically derived objective functions that seek to find a set of image parameters that best explain the measured data. The statistical reconstruction approach can be contrasted to analytic approaches, which seek to invert the mathematical transformations that link the image

and data domains. The most widely used analytic PET image reconstruction method is filtered backprojection (FBP), which is based on the inverse Radon transform (see Section 2.2.2 for details), and its 3D implementations [122–125]. This thesis focuses on statistical PET image reconstruction due to its well-known advantages in accurate system modelling and its flexibility to include a wide range of noise models.

2.2.1 Image Reconstruction as an Inverse Problem

The reconstruction of images from PET data is an inverse problem where there exists an I-dimensional vectorised measured data $\boldsymbol{m}$ (e.g. the measured sinogram) from which the spatio-temporal distribution of a set of parameters is to be estimated. Although the physical object in the scanner consists of a continuous spatio-temporal distribution of parameters, it is usual to discretise the image reconstruction problem to estimate a J-dimensional parameter vector $\boldsymbol{\theta}$ that is linked to the continuous distribution as detailed below. In general, a model that links these discretised parameters of interest to the data/sinogram domain is known *a priori* such that

$$\boldsymbol{q}(\boldsymbol{\theta}) = g(\boldsymbol{\theta}), \tag{2.1}$$

where $\boldsymbol{q}$ is the expected data vector given the input parameters $\boldsymbol{\theta}$ and $g(\cdot)$ is the model between domains. The inverse problem can then be defined with the aim of finding the optimal set of parameters $\hat{\boldsymbol{\theta}}$, such that

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} d(\boldsymbol{m}, \boldsymbol{q}(\boldsymbol{\theta})), \tag{2.2}$$

where $d(\cdot, \cdot)$ is some measure of the distance or divergence between two vectors. The problem defined by Equations 2.1 and 2.2 represents the most general inverse problem, allowing $\boldsymbol{q}$ to be non-linear in $\boldsymbol{\theta}$.

In PET imaging, the model $g(\cdot)$ can be split into two components: a tomographic component $h(\cdot)$ and a component that maps the parameters of interest into a spatial distribution of radioactivity concentration at a given time. To

keep the parameters of interest general, the spatio-temporally continuous distribution of radioactivity concentration, $f(x, y, z, t)$, is denoted as a function of the parameters of interest such that $f(\boldsymbol{\theta}) \stackrel{\text{def}}{=} f(x, y, z, t; \boldsymbol{\theta})$, where the spatio-temporal coordinates are omitted for clarity.

Given the previous definition, the full model can be written as:

$$\boldsymbol{q} = \mathcal{S}(h(f(\boldsymbol{\theta}))), \tag{2.3}$$

where h is a function that maps between the image and data domains (the tomographic model) and $\mathcal{S}$ is a sampling operator to produce a discretised data vector.

2.2.2 Image Representation and the Tomographic Model

Image Representation In order to solve the image reconstruction problem defined by Equations 2.1 and 2.2, the parameters of interest that are to be estimated must be chosen, and a suitable model between these parameters and the continuous radioactivity distribution defined. The most common choice is for $\boldsymbol{\theta}$ to represent the intensities of a set of equally-sized, non-overlapping voxels that extend throughout the field-of-view, so that

$$f(\boldsymbol{\theta}) = \sum_{j=1}^{J} \theta_j \text{rect}(x - x_j, y - y_j, z - z_j, t - t_j), \tag{2.4}$$

where x_j, y_j, z_j and t_j are the spatio-temporal coordinates of the voxel whose intensity is θ_j, J is the number of voxels in the image, and $\text{rect}(x, y, z, t)$ is a rectangular function of fixed spatio-temporal extent, selected to provide non-overlapping, abutting voxels.

Of course, one can use any model linking the parameters of interest to the continuous radioactive distribution. The fundamental characteristic of any such model is whether or not the model is linear in θ_j. When this is the case, the set of functions that are multiplied by $\boldsymbol{\theta}$ can be referred to as 'basis functions'. Equation 2.4 represents a linear model where the parameters are multiplied by a spatially invariant basis function, but in general one can define

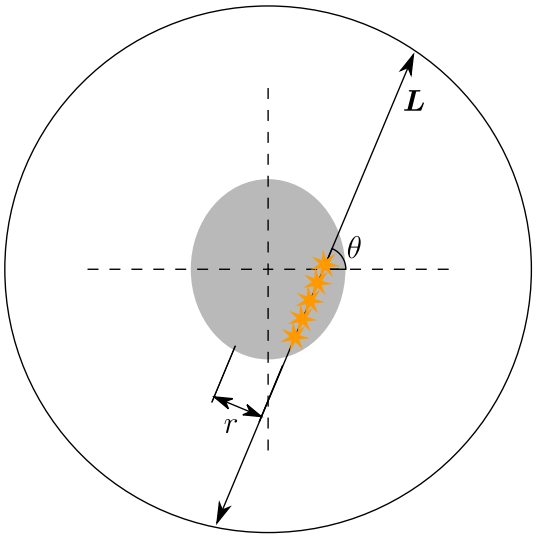


Figure 2.5: Applicability of the Radon transform to 2D PET imaging. The count rate observed for the LOR, $\boldsymbol{L}$, characterised by r, θ, is given by the line integral of recorded counts produced by each of the displayed radioactive events. Each one of these events contributes an equal amount to the total value of $m(r, \theta)$, ignoring the effects of attenuation and finite detector size.

a separate basis function for each of the parameters to be estimated, i.e.:

$$f(\boldsymbol{\theta}) = \sum_{j=1}^{J} \theta_j B_j(x, y, z, t). \tag{2.5}$$

The choice of basis functions can affect the convergence of reconstruction algorithms depending on their orthogonality. Note that the voxel basis functions in Equation 2.4 are orthogonal.

Non-linear models between $\boldsymbol{\theta}$ and $f(\boldsymbol{\theta})$ are less common, and are generally only used in dynamic reconstruction contexts due to the inherent non-linearity between parameters such as binding potential and image intensity [126]. This thesis considers only linear models, due to both their simplicity for implementation, and for their widespread applicability to static PET image reconstruction. Furthermore, within this work all image models are based on spatially invariant, isotropic voxels as demonstrated by Equation 2.4, unless otherwise noted.

The Tomographic Model Once the relationship between the parameters of interest and the spatio-temporal radioactivity distribution is established, the

tomographic model, $h(\cdot)$, must also be defined. The foundation for the 2D tomographic model is the Radon transform, which relates the spatial distribution of an observable quantity with the 'projections' measured as integrals of that quantity in sets of parallel lines [127]. The applicability of the Radon transform in 2D PET imaging is demonstrated in Figure 2.5. It can be seen that any β^+-emitting decays that occur along a given LOR contribute equally to the total counts measured in the corresponding sinogram bin (ignoring resolution degrading effects). This is in agreement with the Radon transform, which can be expressed as:

$$q(r, \theta) = \int_{\boldsymbol{L}} f(x, y) \mathrm{d}\boldsymbol{L}, \tag{2.6}$$

where $\boldsymbol{L}$ denotes the LOR indicated by a particular r, θ pairing and $q(r, \theta)$ is the continuous version of the expected data. An analogous transform exists for the 3D case, known as the 3D x-ray transform [124].

When considering discrete expected sinogram data, $\boldsymbol{q}$, and a voxel-intensity parameter model as given by Equation 2.4, the x-ray transform (in any number of dimensions) can be re-written as a matrix-vector multiplication,

$$q_i = \sum_{j=1}^{J} X_{ij} \theta_j, \tag{2.7}$$

where $\boldsymbol{X}$ is the discrete x-ray transform matrix in which each element X_{ij} denotes the proportional contribution to the counts in sinogram bin i from image voxel j. The simplest model is to set $X_{ij} = 1$ for all voxels which intersect the LOR at all. More advanced methods that take account of the size of the intersection are more typical in PET modelling, such as Siddon's algorithm [128].

In general though, the x-ray transform alone is not sufficient to fully model the PET acquisition process due to photon attenuation and spatially varying sensitivity. These effects can be included in the tomographic model relatively

easily [129] by rewriting Equation 2.7 with additional terms as

$$q_i = \sum_{j=1}^{J} A_i N_i X_{ij} \theta_j. \tag{2.8}$$

Here A_i represents the fraction of unattenuated photon pairs that reach the detector rings along LOR i, and N_i represents the proportion of incident photon pairs that are detected in sinogram bin i.

As described above, other effects alter the number of counts recorded in each sinogram bin, such as randoms, scatters, and resolution degradation. These can also be included in the tomographic model, with randoms and scatters modelled typically as an additive background term, $\boldsymbol{b}$, that is independent of $\boldsymbol{\theta}$:

$$q_i = \sum_{j=1}^{J} A_i N_i X_{ij} \theta_j + b_i. \tag{2.9}$$

To model the various resolution degrading effects listed in Section 2.1.1, each one may be introduced separately into the model in the appropriate position. For example, positron range occurs in image space and inter-detector scatter occurs in sinogram space [118]. Nonetheless, it is common to aggregate all of these resolution degrading effects into a single equivalent image space PSF model [129], according to:

$$q_i = \sum_{j=1}^{J} A_i N_i X_{ij} \sum_{l=1}^{J} H_{jl} \theta_l + b_i. \tag{2.10}$$

where the matrix $\boldsymbol{H}$ contains the (potentially spatially variant) PSF model.

Overall, whatever effects are included, the tomographic model of PET data arising from a discrete, spatially isotropic voxel grid of intensities $\boldsymbol{\theta}$ can be written as the affine relation:

$$q_i = \sum_{j=1}^{J} P_{ij} \theta_j + b_i. \tag{2.11}$$

Here P_{ij} denotes the total probability of a radioactive decay in voxel j being detected in the detector pair i, including the effects of scanner geometry, attenuation, normalisation, and resolution degradation. This matrix $\boldsymbol{P}$ that

includes all of the required effects is known as the system matrix.

The linearity of this model in θ_j is important in the following sections where a derivation of a reconstruction algorithm is considered. It is important to note that this linearity is in fact an assumption; at high counts levels, due to dead-time effects and the finite coincidence timing window τ, the output q_i stops being linear in θ_j, and at higher count levels the scanner system can fully saturate [130].

2.2.3 The Maximum Likelihood Objective Function

As shown above, PET data acquisition is fundamentally a counting process; the number of coincidences observed during the scan duration for sinogram bin i is what constitutes the measured data m_i. If the counts observed in each bin are assumed to be independent variables, m_i can be described as Poisson-distributed according to:

$$m_i \sim \text{Poisson}(\lambda_i), \tag{2.12}$$

where λ_i is the underlying mean number of counts per bin.

Noting the Poisson-distributed nature of measured PET data allows the definition of the likelihood of an image estimate $\boldsymbol{\theta}$ giving rise to the observed data $\boldsymbol{m}$. Although alternative distributions have been proposed in the literature (e.g. [131–133]), this thesis considers only the Poisson noise model. Nonetheless, the methods presented in this thesis can be extended to the case of alternative noise models.

Let $\Pr(m_i|q_i)$ denote the probability of observing m_i coincidences in sinogram bin i when the expected number of counts is q_i. Since m_i is a Poisson-distributed variable, $\Pr(m_i|q_i)$ is then given by:

$$\Pr(m_i|q_i) = \frac{e^{-q_i} q_i^{m_i}}{m_i!}. \tag{2.13}$$

Under the assumption that the observation of counts in each bin is an independent random process, the total probability of observing $\boldsymbol{m}$ given $\boldsymbol{q}$ is

the product of the probability of observing each bin's count level separately, i.e:

$$\Pr(\boldsymbol{m}|\boldsymbol{q}) = \prod_{i=1}^{I} \frac{e^{-q_i} q_i^{m_i}}{m_i!}. \tag{2.14}$$

Given that the expected data $\boldsymbol{q}$ is a function of the parameters of interest $\boldsymbol{\theta}$ according to Equation 2.1, $\boldsymbol{\theta}$ can be substituted into Equation 2.14 to give $\Pr(\boldsymbol{m}|\boldsymbol{\theta})$. The likelihood of a parameter vector $\boldsymbol{\theta}$ giving rise to the measured data, $L(\boldsymbol{\theta}|\boldsymbol{m})$, is then defined as equal to $\Pr(\boldsymbol{m}|\boldsymbol{\theta})$.

When considering the standard, spatially invariant voxel image representation given in Equation 2.4, the linear tomographic model given in Equation 2.11 can be used to relate $\boldsymbol{\theta}$ to $\boldsymbol{q}$. This leads to the following expression for the likelihood of a parameter set, $\boldsymbol{\theta}$, giving rise to the measured data $\boldsymbol{m}$:

$$L(\boldsymbol{\theta}|\boldsymbol{m}) = \prod_{i=1}^{I} \frac{\exp\left(-\sum\limits_{j=1}^{J} P_{ij}\theta_j - b_i\right)\left(\sum\limits_{j=1}^{J} P_{ij}\theta_j + b_i\right)^{m_i}}{m_i!}. \tag{2.15}$$

Maximum likelihood (ML) reconstruction methods aim to find the parameters $\boldsymbol{\theta}$ that maximise this likelihood, thereby providing image estimates that best explain the measured data. In practice, the logarithm of Equation 2.15 is taken, in order to simplify the derivation of iterative reconstruction algorithms. This yields the Poisson log-likelihood objective function for ML reconstruction, $\Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m})$, which forms the foundation of clinically used PET image reconstruction algorithms, given by:

$$\begin{aligned}\Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m}) &= \ln(L(\boldsymbol{\theta}|\boldsymbol{m})) \\ &= \sum_{i=1}^{I}\left[m_i \ln\left(\sum_{j=1}^{J} P_{ij}\theta_j + b_i\right) - \sum_{j=1}^{J} P_{ij}\theta_j - b_i - \ln(m_i!)\right].\end{aligned} \tag{2.16}$$

The set of parameters that maximise this objective function together make up the ML solution, denoted $\boldsymbol{\theta}^{\mathrm{ML}}$. Note that this objective function is concave and equivalent to the negative Kullback-Leibler divergence [134]. Methods for estimating $\boldsymbol{\theta}^{\mathrm{ML}}$ are discussed in Section 2.3.

2.2.4 Maximum *A Posteriori* and Penalised Maximum Likelihood Objective Functions

As opposed to the maximum likelihood objective function defined in Section 2.2.3, an alternative objective function can be defined using Bayesian statistics. According to the well-known Bayes' Theorem, the posterior probability of $\boldsymbol{\theta}$ given $\boldsymbol{m}$ is given by:

$$\Pr(\boldsymbol{\theta}|\boldsymbol{m}) = \frac{\Pr(\boldsymbol{m}|\boldsymbol{\theta})\Pr(\boldsymbol{\theta})}{\Pr(\boldsymbol{m})}, \tag{2.17}$$

where $\Pr(\boldsymbol{\theta})$ is the probability of the scanned object corresponding to the parameter vector $\boldsymbol{\theta}$, and $\Pr(\boldsymbol{m})$ is the probability of a measured data vector $\boldsymbol{m}$ occurring. Since $\boldsymbol{m}$ are observed outcomes, $\Pr(\boldsymbol{m}) = 1$. Considering a known (or assumed) probability distribution of $\boldsymbol{\theta}$, known as the prior, the posterior $\Pr(\boldsymbol{\theta}|\boldsymbol{m})$ is therefore given by:

$$\Pr(\boldsymbol{\theta}|\boldsymbol{m}) = \Pr(\boldsymbol{m}|\boldsymbol{\theta})\Pr(\boldsymbol{\theta}). \tag{2.18}$$

Assuming that $\Pr(\boldsymbol{\theta})$ takes the form

$$\Pr(\boldsymbol{\theta}) = Ce^{-\beta U(\boldsymbol{\theta})}, \tag{2.19}$$

where $U(\boldsymbol{\theta})$ is some function of $\boldsymbol{\theta}$ and C is a normalising factor to bring the sum of the probabilities to 1, the posterior probability becomes:

$$\Pr(\boldsymbol{\theta}|\boldsymbol{m}) = Ce^{-\beta U(\boldsymbol{\theta})} \prod_{i=1}^{I} \frac{\exp\left(-\sum_{j=1}^{J} P_{ij}\theta_j - b_i\right)\left(\sum_{j=1}^{J} P_{ij}\theta_j + b_i\right)^{m_i}}{m_i!}. \tag{2.20}$$

Taking the logarithm of this posterior yields the log-posterior objective

function, used for maximum *a posteriori* (MAP) PET image reconstruction:

$$\Phi^{\mathrm{MAP}}(\boldsymbol{\theta}|\boldsymbol{m}) = \sum_{i=1}^{I}\left[m_i \ln\left(\sum_{j=1}^{J} P_{ij}\theta_j + b_i\right) - \sum_{j=1}^{J} P_{ij}\theta_j - b_i - \ln(m_i!)\right] - \beta U(\boldsymbol{\theta}) + C. \tag{2.21}$$

Ignoring the constant C, this is equivalent to:

$$\Phi^{\mathrm{MAP}}(\boldsymbol{\theta}|\boldsymbol{m}) = \Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m}) - \beta U(\boldsymbol{\theta}). \tag{2.22}$$

In this form, it is clear that the MAP objective function can be considered a penalised maximum likelihood (PML) objective function resulting from the addition of a general penalty term $U(\boldsymbol{\theta})$ onto Equation 2.16. Whilst strictly speaking priors are a subset of penalties insofar as not all penalties will necessarily satisfy the conditions required to define a true prior distribution of $\boldsymbol{\theta}$, the PET image reconstruction literature tends to use the terms 'maximum *a posteriori*' and 'penalised maximum likelihood', and 'prior' and 'penalty' interchangeably. The two terms are used interchangeably in this thesis to reflect this common usage. Note that intuitively speaking, the ML case presented in Section 2.2.3 corresponds to the MAP case where all images $\boldsymbol{\theta}$ are *a priori* considered equally probable.

One particularly important family of penalties for MAP image reconstruction in PET is the Markov random field (MRF) penalties. These penalties encourage local smoothness by ensuring that neighbouring voxels do not differ too greatly in their intensity. In general, an MRF penalty has the following form

$$U(\boldsymbol{\theta}) = \sum_{j=1}^{J}\sum_{k\in\mathcal{N}_j} w_{jk}\psi(\theta_j, \theta_k), \tag{2.23}$$

where $\mathcal{N}_j$ is the set of voxel indices that make up a neighbourhood around voxel j, w_{jk} are a set of coupling weights between voxel j and each of its neighbours k, and $\psi(\theta_j, \theta_k)$ is usually a symmetric function that provides a high penalty contribution for neighbouring voxel pairs that are expected to be similar in intensity.

Typical examples of MRF penalties include the quadratic penalty, where $\psi_{\text{quad}}(\theta_j, \theta_k) = (\theta_j - \theta_k)^2$ [135] or ℓ_1-based total variational (TV) penalties, such as $\psi_{\text{TV}}(\theta_j, \theta_k) = |\theta_j - \theta_k|$. This latter penalty has received attention in the context of PET image reconstruction due to its favourable edge-preserving properties [136–140]. However, the strict TV penalty is not continuously differentiable, which complicates the solving of the optimisation problem. For this reason, smooth versions of the TV penalty have been proposed, including the Huber prior [141], the Lange prior [142], the log-cosh prior [143], and the hyperbolic prior [136, 144]. In general these methods provide a $\psi(\theta_j, \theta_k)$ that is proportional to $(\theta_j - \theta_k)^2$ for $\theta_j - \theta_k < \varepsilon$ and proportional to $|\theta_j - \theta_k|$ for $\theta_j - \theta_k > \varepsilon$, where ε is an arbitrary cut-off point.

2.3 Solution of the Statistical PET Image Reconstruction Problem

Once the objective function has been defined, a method to find the values of $\boldsymbol{\theta}$ that optimise it must be selected. The most common PET image reconstruction algorithms are based on the principle of expectation-maximisation (EM). Compared to other methods, EM algorithms as used in PET image reconstruction are easier to implement, are more well-adapted to the Poisson log-likelihood objective function, and automatically enforce the non-negativity constraint that exists in PET imaging. For these reasons, this work focuses on EM-based reconstruction methods. Nonetheless, reflecting the wide range of algorithms available in the field of optimisation, there are a variety of methods that have been applied to PET image reconstruction, including conjugate gradient methods [145], iterative coordinate descent [146] and more recently more advanced methods such as primal-dual [147], Bregman [148], and alternating direction [139, 140] methods, and hybrid methods defined specifically for PET image reconstruction (e.g. [137]).

2.3.1 Expectation-Maximisation Algorithms

The maximum likelihood estimate of $\boldsymbol{\theta}$, denoted $\boldsymbol{\theta}^{\mathrm{ML}}$, is the set of parameters for which the objective function $\Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m})$ is at a maximum, i.e. at which the gradient of $\Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m})$ is equal to zero. This can be written as

$$\boldsymbol{\theta}^{\mathrm{ML}} = \left\{ \hat{\theta}_j : \left. \frac{\partial \Phi^{\mathrm{ML}}}{\partial \theta_j} \right|_{\hat{\boldsymbol{\theta}}} = 0 \; \forall j \right\}, \tag{2.24}$$

where $\frac{\partial \Phi^{\mathrm{ML}}}{\partial \theta_j}$ is the partial derivative of the objective function with respect to the jth element of $\boldsymbol{\theta}$. Considering Equation 2.16, this gradient is given by

$$\frac{\partial \Phi^{\mathrm{ML}}}{\partial \theta_j} = \sum_{i=1}^{I} \frac{m_i P_{ij}}{\sum\limits_{l=1}^{J} P_{il}\theta_l + b_i} - \sum_{i=1}^{I} P_{ij}. \tag{2.25}$$

Since setting this equation equal to zero and rearranging to give a closed-form expression for θ_j is not possible, iterative methods are employed to find $\boldsymbol{\theta}^{\mathrm{ML}}$.

The most common iterative algorithm for finding $\boldsymbol{\theta}^{\mathrm{ML}}$ is the EM algorithm, which was proposed by Shepp and Vardi [149] and Lange and Carson [150] in the 1980s. The resulting algorithm is known as maximum likelihood EM, or MLEM. The EM algorithm can be defined or derived in multiple ways, but in the most common interpretation it uses an unknown set of 'complete data', z_{ij}, representing the number of emissions from each voxel j detected in each sinogram bin i. This allows the definition of a two-step process for updating the voxel intensities as follows:

In step 1 the *expectation* of the complete data, $\bar{z}_{ij}$, is calculated, based on both the measured data $\boldsymbol{m}$ and the current image estimate $\boldsymbol{\theta}^{(\nu)}$, according to:

$$\bar{z}_{ij} = \frac{P_{ij}\theta_j^{(\nu)}}{\sum\limits_{l=1}^{J} P_{il}\theta_l^{(\nu)} + b_i} m_i. \tag{2.26}$$

Explicitly, it is assumed that both $\boldsymbol{m}$ and $\boldsymbol{\theta}^{(\nu)}$ are known, and the value of $\bar{\boldsymbol{z}}$ that is consistent with both is calculated.

In step 2, the complete data version of the Poisson log-likelihood is *maximised* according to:

$$\boldsymbol{\theta}^{(\nu+1)} = \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\bar{z}_{ij} \ln(P_{ij}\theta_j) - P_{ij}\theta_j - \ln(\bar{z}_{ij}!)\right]. \tag{2.27}$$

In contrast to directly maximising Equation 2.15, the solution of Equation 2.27 is possible due to the introduction of the complete data, which ensures mathematical tractability. By differentiating and setting to zero, the updated image, $\boldsymbol{\theta}^{(\nu+1)}$, can be found according to:

$$\theta_j^{(\nu+1)} = \frac{\sum\limits_{i=1}^{I} \bar{z}_{ij}}{\sum\limits_{i=1}^{I} P_{ij}}. \tag{2.28}$$

Substituting Equation 2.26 into Equation 2.28 results in the overall MLEM update formula:

$$\theta_j^{(\nu+1)} = \frac{\theta_j^{(\nu)}}{\sum\limits_{i=1}^{I} P_{ij}} \sum_{i=1}^{I} P_{ij} \frac{m_i}{\sum\limits_{l=1}^{J} P_{il}\theta_l^{(\nu)} + b_i}. \tag{2.29}$$

Note that the MLEM update as usually implemented is multiplicative, in contrast to the additive gradient-based methods more typical in iterative optimisation algorithms, although MLEM is expressible as a scaled gradient ascent method [134]. MLEM is guaranteed to converge to the ML estimate given enough iterations, and enforces physically possible non-negative images if initialised with a non-negative image.

EM can also be employed to solve the PML/MAP problem, which also does not allow for a direct solution. The expectation step of EM for MAP problems (MAPEM) is identical to that given in Equation 2.26, but the maximisation step is altered by the existence of the penalty term in the objective function. Concretely, the surrogate objective function becomes:

$$\boldsymbol{\theta}^{(\nu+1)} = \arg\max_{\boldsymbol{\theta}} \left\{ \sum_{i=1}^{I} \sum_{j=1}^{J} \left[\bar{z}_{ij} \ln(P_{ij}\theta_j) - P_{ij}\theta_j - \ln(\bar{z}_{ij}!)\right] - \beta U(\boldsymbol{\theta}) \right\}. \tag{2.30}$$

Differentiating this equation and setting to zero yields the following expression:

$$\frac{1}{\theta_j}\sum_{i=1}^{I}\bar{z}_{ij} - \sum_{i=1}^{I}P_{ij} - \beta\frac{\partial U}{\partial\theta_j} = 0 \tag{2.31}$$

The image update is then the value of $\boldsymbol{\theta}$ that satisfies Equation 2.31. It is clear that the ability to find a solution depends on the form of $U(\boldsymbol{\theta})$. For some choices of $U(\boldsymbol{\theta})$, Equation 2.31 is directly solvable, as it is for the MLEM case described above. In the case of MRF-priors as described above, the spatial relationships between voxels generally renders Equation 2.31 intractable.

One method of dealing with the mathematical intractability of Equation 2.31 for differentiable penalties is the 'one-step-late' (OSL) approach proposed by Green [143, 151], whereby it is assumed that the gradient at the unknown updated image estimate is approximately equal to the gradient at the current image estimate, i.e.

$$\left.\frac{\partial U}{\partial\theta_j}\right|_{\boldsymbol{\theta}^{(\nu+1)}} \approx \left.\frac{\partial U}{\partial\theta_j}\right|_{\boldsymbol{\theta}^{(\nu)}}. \tag{2.32}$$

Using this approximation, Equation 2.31 can be rearranged to yield the updated image $\boldsymbol{\theta}^{(\nu+1)}$:

$$\theta_j^{(\nu+1)} = \frac{\theta_j^{(\nu)}}{\sum_{i=1}^{I}P_{ij} + \beta\left.\frac{\partial U}{\partial\theta_j}\right|_{\boldsymbol{\theta}^{(\nu)}}}\sum_{i=1}^{I}P_{ij}\frac{m_i}{\sum_{l=1}^{J}P_{il}\theta_l^{(\nu)} + b_i}. \tag{2.33}$$

As can be seen, this update formula is similar to the MLEM update in Equation 2.29, and so is simple to implement as long as the gradient of the penalty is known. Due to its simplicity, it is one of the most common algorithms used for penalised PET image reconstruction and is included in STIR (Software for Tomographic Image Reconstruction), one of the most popular open-source PET image reconstruction software packages available [152].

The OSL-MAPEM method does, however, have disadvantages that offset its simplicity. Firstly, the approximation in Equation 2.32 means that the update in 2.33 is not actually solving the maximisation problem of Equation 2.22, which means that OSL-MAPEM is not guaranteed to converge as MLEM is. Secondly, if β is too large, image estimates can become negative using OSL-

MAPEM, an issue that is not possible when using MLEM.

An alternative method for performing MAPEM is to find a surrogate for $U(\boldsymbol{\theta})$ that allows a closed form solution to be found [153]. For quadratic MRF priors of the form

$$U(\boldsymbol{\theta}) = \frac{1}{4}\sum_{j=1}^{J}\sum_{k\in\mathcal{N}_j} w_{jk}\left(\theta_j - \theta_k\right)^2, \tag{2.34}$$

a suitable surrogate is [154]

$$U_s(\boldsymbol{\theta}) = \frac{1}{4}\sum_{j=1}^{J}\sum_{k\in\mathcal{N}_j} w_{jk}\left(2\theta_j - \theta_j^{(\nu)} - \theta_k^{(\nu)}\right)^2. \tag{2.35}$$

Substituting this surrogate penalty in place of $U(\boldsymbol{\theta})$ in Equation 2.30 and solving for $\boldsymbol{\theta}$ gives the following update formula:

$$\theta_j^{(\nu+1)} = \frac{-A_j^{(\nu)} + \sqrt{\left(A_j^{(\nu)}\right)^2 + 4C_jB_j^{(\nu)}}}{2C_j}, \tag{2.36}$$

where

$$A_j^{(\nu)} = \sum_{i=1}^{I} P_{ij} - \beta\sum_{k\in\mathcal{N}_j} w_{jk}\left(\theta_j^{(\nu)} + \theta_k^{(\nu)}\right), \tag{2.37}$$

$$B_j^{(\nu)} = \theta_j^{(\nu)}\sum_{i=1}^{I} P_{ij}\frac{m_i}{\sum_{l=1}^{J} P_{il}\theta_l^{(\nu)} + b_i}, \tag{2.38}$$

and

$$C_j = 2\beta\sum_{k\in\mathcal{N}_j} w_{jk}. \tag{2.39}$$

Note that 2.37 includes an image space smoothing step, and $B_j^{(\nu)}$ is proportional to the MLEM update in Equation 2.29.

2.4 Related Work in Longitudinal and Multi-Dataset Image Reconstruction

The aim of this thesis is to propose methods of sharing information between PET scans during the reconstruction process with a particular focus on oncol-

ogy treatment response studies and sequential multi-tracer PET scans. While these areas remain largely neglected in the PET image reconstruction literature, there are a number of related approaches in multi-dataset image reconstruction that informed the methodologies proposed in this thesis. This Section briefly reviews these methods.

2.4.1 Anatomically Guided PET Image Reconstruction

In ET reconstruction, the idea of improving images by incorporating information from a complementary imaging modality has existed for many years [155]. This area is related to anatomically based partial volume correction [156, 157], but also includes the case where anatomical information is used for noise reduction. Many of the early approaches for anatomically guided ET image reconstruction used segmentations or edge maps of the anatomical images (also known as prior images) in order to define edges over which smoothing should not be applied [135, 158–164]. One of the simplest formulations for this approach uses the weighted quadratic penalty (see Section 2.2.4), given by:

$$U(\boldsymbol{\theta}) = \sum_{j=1}^{J} \sum_{k \in \mathcal{N}_j} w_{jk} \left(\theta_j - \theta_k\right)^2, \tag{2.40}$$

where the weights w_{jk} can be set based on the *a priori* presence of an anatomical edge between voxels j and k [135, 162, 163]. Non-quadratic functionals can also be used in place of the quadratic [160]. An alternative related formulation is to introduce the concept of 'line-sites', a set of binary parameters that define the existence or non-existence of an edge between two voxels [158, 159]. In these methods, the line-site parameters are also to be estimated during the reconstruction, unlike the weighted quadratic penalty, where weights are usually held fixed throughout the reconstruction. Anatomical guidance can be used to constrain the construction of these line-sites in the reconstruction [158, 159].

Beyond using anatomical segmentations to define edge-based penalties, other methods have used segmentations of anatomical images to define region-based statistical models to serve as priors [160, 165–168]. In these priors, a

statistical model is created for each region (e.g. a Gaussian distribution for grey matter voxels), and the penalty term is the deviation of reconstructed image voxel values from their assumed regional distributions. These methods can either perform the PET image segmentation during the reconstruction, with the aid of anatomical segmentations [165], or they can rely entirely on the anatomical segmentations (e.g. [166]). The performance of these methods is dependent on the quality of the anatomical segmentations, the assumed regional intensity distributions and the methods of estimating regional distribution parameters.

A segmentation-free anatomically guided reconstruction was proposed by Bowsher [169], in which the weights for a quadratic penalty (Equation 2.40) were set directly from MR voxel intensities. In the original method, using the MR image, the B voxels with the closest intensities to voxel j were selected from a local search window. The corresponding weights w_{jk} were then set to 1 to encourage smoothing, and all other w_{jk} were set to 0. This method of setting weights removes the need for segmentation or edge-detection of anatomical images, and has been used and modified in various guises in the anatomically guided ET reconstruction literature [147, 168, 170–172].

Another option for segmentation-free anatomical guidance using local image structure is to promote joint edges in the ET and anatomical images [144, 147, 173], for example by using the parallel level sets approach [147, 173]. In a rather unique work, Ardekani *et al* [134] performed inter-iteration MR-guided filtering, before using the resultant image as a prior in a manner reminiscent of the region-based statistical priors mentioned above.

An alternative paradigm for segmentation-free anatomically guided ET image reconstruction is to use information theoretic priors such as joint entropy and mutual information [168, 174–179]. In these cases, the prior is defined in terms of the joint probability density function (PDF) of the ET image and the anatomical image. These methods have the advantage of acting directly on the anatomical image voxel values, requiring no processing or calculation of weights. On the other hand, when used alone, mutual information and joint entropy priors do not acknowledge the spatial relationships between voxels, i.e.

any reordering of voxels in both the ET and the anatomical image will yield the same penalty value. This can be addressed by including a spatial prior alongside the information theoretic prior [176], or by using feature vectors designed to reflect spatial structure in the prior calculation [175, 178, 179].

All of the above-mentioned methods apply anatomical guidance in a PML or MAP framework. A different approach to including side information into PET image reconstruction problems, known as the kernel method or kernelised EM (KEM), was recently proposed [180] and has since been widely investigated as a method for including MR information in PET image reconstruction [181–186]. Despite its widely used name, KEM is actually the standard MLEM algorithm applied to a re-parameterised image reconstruction problem defined by using the kernel trick from machine learning [180]. The KEM method can be defined simply. First, a so-called 'kernel matrix' is calculated according to

$$K_{jk} = \begin{cases} f_K(\boldsymbol{f}_j^{\text{prior}}, \boldsymbol{f}_k^{\text{prior}}), & k \in \mathcal{S}_j \\ 0, & \text{otherwise} \end{cases} \tag{2.41}$$

where $f_K(\cdot, \cdot)$ is a kernel function which effectively quantifies the similarity between two vectors; $\boldsymbol{f}_j^{\text{prior}}$ is a feature vector extracted from the source of prior information, corresponding to voxel j in the target ET image; and $\mathcal{S}_j$ is some subset of voxels in the ET image. A common choice for f_K is the radial Gaussian function given by [180]:

$$f_K(\boldsymbol{f}_j^{\text{prior}}, \boldsymbol{f}_k^{\text{prior}}) = \exp\left(-\frac{\left\|\boldsymbol{f}_j^{\text{prior}} - \boldsymbol{f}_k^{\text{prior}}\right\|_2^2}{2\sigma^2}\right), \tag{2.42}$$

where σ is a scaling parameter that can also be used to adjust the strength of similarity values. The PET image to be estimated is then modelled as:

$$\boldsymbol{\theta} = \boldsymbol{K}\boldsymbol{\alpha}, \tag{2.43}$$

where the vector $\boldsymbol{\alpha}$ is a set of coefficients that control the contribution to the final image $\boldsymbol{\theta}$ from each column of the kernel matrix. In this sense, the

kernel method is a basis function approach that groups similar voxels together in the reconstruction based on their similarity in the prior image. Since $\boldsymbol{\theta}$ is linear in $\boldsymbol{\alpha}$, and the measured data $\boldsymbol{q}$ is linear in $\boldsymbol{\theta}$, the overall system model (c.f. Equation 2.11) is linear in the parameters of interest. Therefore, the coefficients $\boldsymbol{\alpha}$ can be estimated using the MLEM algorithm [180] with only minor modifications.

The inclusion of the subsets $\mathcal{S}_j$ in the kernel matrix calculation in Equation 2.41 performs sparsification on the kernel matrix in order to 1) reduce the computational burden of the method, and 2) reduce the effect of inaccurate similarity measurements from suboptimal choices of features or kernel function. Some common examples of sparsification subsets include k-nearest neighbours, local image neighbourhoods, or the so-called ε-ball method [180], respectively given by:

$$\mathcal{S}_j = \left\{ k : \boldsymbol{f}_k^{\text{prior}} \in \text{kNN}\left(\boldsymbol{f}_j^{\text{prior}}\right) \right\} \tag{2.44a}$$

$$\mathcal{S}_j = \left\{ k \in \mathcal{N}_j \right\} \tag{2.44b}$$

$$\mathcal{S}_j = \left\{ k : d(\boldsymbol{f}_j^{\text{prior}}, \boldsymbol{f}_k^{\text{prior}}) < \varepsilon \right\}, \tag{2.44c}$$

where $\text{kNN}\left(\boldsymbol{f}_j^{\text{prior}}\right)$ is the set of k-nearest neighbours to $\boldsymbol{f}_j^{\text{prior}}$ in feature space, $d(\cdot,\cdot)$ is any distance function, and ε is a distance threshold. Note that Equations 2.44a and 2.44b give a fixed number of non-zero kernel matrix values for each voxel, whereas Equation 2.44c gives an adaptive number of non-zero kernel elements.

Another area in which the dichotomy between PML/MAP methods and re-parameterisation is apparent is in the reconstruction of PET images using anatomically derived dictionaries [187, 188]. In the first case, Chen *et al* [187] used a MAP approach to incorporate a CT-based patch dictionary, penalising images that could not be expressed as a sparse collection of dictionary patches. On the other hand, Tahaei and Reader [188] explicitly modelled PET images as a weighted combination of their dictionary patches, regularising the results with a sparsity constraint. In general, it remains an open question as to

whether MAP-based methods or re-parameterisation approaches are superior, and in which situations.

An alternative to including anatomical information during the reconstruction at all is to include the information post-reconstruction. In theory, it is possible to include almost any reconstruction prior in a post-reconstruction approach by replacing $L(\boldsymbol{\theta}|\boldsymbol{m})$ with $L(\boldsymbol{\theta}|\boldsymbol{\theta}^{\mathrm{ML}})$, i.e. by replacing the measured data with the ML image estimate in the log-likelihood (and adapting the likelihood function accordingly) [189]. Including priors in such a post-reconstruction approach has the advantage of lower computational burden, but is suboptimal unless the noise is correctly accounted for [182, 189]. Nonetheless, including anatomical information via post-processing has received interest in recent years, especially with the introduction of the non-local means (NLM) filter. The NLM filter is a shift-variant filter that was introduced in the context of natural image processing, where it was defined as follows [190]:

$$\theta_j^{\mathrm{NLM}} = \sum_{k=1}^{J} w_{jk}^{\mathrm{NLM}} \theta_k, \tag{2.45}$$

with

$$w_{jk}^{\mathrm{NLM}} = C \exp\left(-\frac{\|\boldsymbol{f}_j - \boldsymbol{f}_k\|_2^2}{2\sigma^2}\right), \tag{2.46}$$

where C is a normalising factor, σ is a shape parameter, and $\boldsymbol{f}_j$ represents a patch centred on voxel j, rearranged as a vector. In essence, the NLM filter in this form quantifies the similarity between two pixels in terms of the Euclidean distance between their local patch vectors, and uses these similarities as spatially variant weights in a weighted average filter. In its application to the denoising of natural images, patches are used to alleviate the effects of noise and to consider local structure when calculating similarity. While this original implementation of NLM filtering has applications in ET [191], an important possibility raised is in incorporating anatomical information into a NLM filtering in PET by including a NLM filter derived from the anatomical image. One method to do this is to modulate a PET-based NLM filter by a binary value based on voxel variation in the anatomical image [192]. A more

general method is to define the anatomically guided NLM (A-NLM) filter given by

$$w_{jk}^{\text{A-NLM}} = C\exp\left(-\frac{\left\|\boldsymbol{f}_j^{\text{prior}} - \boldsymbol{f}_k^{\text{prior}}\right\|_2^2}{2\sigma^2}\right), \tag{2.47}$$

where $\boldsymbol{f}_j^{\text{prior}}$ is the patch extracted from voxel j in the anatomical image. This filter can be used on its own (as in [180] in a dynamic PET context), or combined with the original ET-based NLM filter to provide a single spatially variant filter for post-reconstruction denoising of ET images [193].

It is worth noting that there are strong similarities between the kernel method, NLM filtering with side information, and Bowsher or other weighted quadratic penalties. Indeed, Wang and Qi [180] used their kernel matrix directly as a guiding NLM filter, which is possible since both matrices essentially contain the same information about the similarity between voxels in the guiding image. Indeed, as originally presented in Equations 2.42 and 2.47, the two are defined with the same kernel; the only difference being the sparsification step included in kernel EM that can easily be extended to the NLM filter. Furthermore, the relationship between NLM filters and weights in a guided MAP reconstruction has also been noted, with Nguyen *et al* referring to their weighted quadratic method as 'non-local regularisation' [170].

2.4.2 Prior Image Constrained Image Reconstruction for X-ray CT

Prior images have also been incorporated into x-ray CT reconstruction through various techniques in recent years. One of the first such methods was the prior image constrained compressed sensing (PICCS) technique [194], which attempts to find domains in which the target image is sparse to enable improved undersampled image recovery according to the principles of compressed sensing [195, 196]. In its originally proposed form, PICCS is formulated as a

constrained optimisation problem defined by

$$\begin{aligned} \boldsymbol{\theta}^{\text{PICCS}} = &\arg\min_{\boldsymbol{\theta}} \left[\alpha \left\| \boldsymbol{\Psi}_1(\boldsymbol{\theta} - \boldsymbol{\theta}^{\text{prior}}) \right\|_1 + (1-\alpha) \left\| \boldsymbol{\Psi}_2 \boldsymbol{\theta} \right\|_1 \right], \\ &\text{s. t. } \boldsymbol{P}\boldsymbol{\theta} = \boldsymbol{m}, \end{aligned} \tag{2.48}$$

where $\boldsymbol{\theta}^{\text{prior}}$ denotes the prior image, $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ denote sparsifying transforms, and $\|\cdot\|_p$ denotes the ℓ_p-norm of a vector. The original proposal of PICCS used the spatial gradient operator for both $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$. Using the hyperparameter α, the PICCS objective function trades off the degree to which $\boldsymbol{\theta}$ agrees with the prior image by having a sparse representation of the difference between the two, and the degree to which $\boldsymbol{\theta}$ itself has sparse representation as utilised in typical compressed sensing methods. PICCS can also be expressed in an unconstrained form, which allows a range of iterative solvers to be used for improved reconstruction performance [197]. Furthermore, non-convex alternatives to the ℓ_1-norm have been shown to allow greater levels of undersampling and faster reconstruction convergence in CT image reconstruction [198].

PICCS has generally been applied in the context of dynamic CT image reconstruction [194, 197, 198], but has also been adapted to dynamic PET image reconstruction [180]. In dynamic imaging the prior image can be formed using early time-point or pre-contrast images [198], or by producing composite images using the entirety of the dynamic data [180, 194, 197].

Similar methods to PICCS have also been applied in the context of longitudinal CT imaging, and these methods have been referred to as prior image penalised likelihood estimation/estimates (PIPLE) [199]. While the basic principle of PIPLE is similar to PICCS, one important advance in the proposal of the PIPLE reconstruction methods is the upgrade from the simple noise model of PICCS to a more accurate Poisson-derived likelihood-based objective function. Furthermore, PIPLE was combined with a registration step to produce the prior-image-registered penalised likelihood estimation (PIRPLE) algorithm that combines the estimation of follow up voxel intensity parameters with estimation of registration parameters between the target image and the

prior image [199]. The PIRPLE problem is then given by:

$$\{\boldsymbol{\theta}^{\text{PIRPLE}}, \boldsymbol{\rho}^{\text{PIRPLE}}\} = \arg\max_{\boldsymbol{\theta},\boldsymbol{\rho}} \Big[\Phi^{\text{ML}}(\boldsymbol{\theta}|\boldsymbol{m}) - \beta_1 \left\|\boldsymbol{\Psi}_1(\boldsymbol{\theta} - \boldsymbol{M}(\boldsymbol{\rho})\boldsymbol{\theta}^{\text{prior}})\right\|_1 - \beta_2 \left\|\boldsymbol{\Psi}_2\boldsymbol{\theta}\right\|_1\Big], \tag{2.49}$$

where β_1 and β_2 are parameters controlling the relative strength of the penalties on the difference image and the target image respectively, and $\boldsymbol{M}(\boldsymbol{\rho})$ is the prior-to-target alignment operator parameterised by the vector $\boldsymbol{\rho}$. This formulation allows for the estimation of rigid alignment [199] or non-rigid alignment [200] simultaneously to the estimation of the follow-up target image. To achieve this, alternating optimisation approaches have been used that update the image estimate and the alignment parameters successively. PIRPLE was originally utilised to improve follow-up images which contained small lesions not present in the prior image. For this reason, the sparsifying transform $\boldsymbol{\Psi}_1$ was selected to be the identity matrix, thereby encouraging sparsity in the difference image directly [199, 200].

Interest has also been paid in the CT image reconstruction literature to so-called reconstruction-of-difference (RoD) methods. These methods are similar to other prior image constrained methods, but reformulate the reconstruction problem so that the voxel-wise intensity differences, $\boldsymbol{\delta}$, between the two images become the parameters that are to be estimated [201–203]. In this case, the CT RoD objective function (ignoring misregistration for simplicity) can be written as:

$$\{\boldsymbol{\delta}^{\text{RoD}}\} = \arg\max_{\boldsymbol{\delta}} \left[\Phi^{\text{ML}}(\boldsymbol{\delta}|\boldsymbol{m}, \boldsymbol{\theta}^{\text{prior}}) - \beta_1 \left\|\boldsymbol{\delta}\right\|_1 - \beta_2 \left\|\boldsymbol{\Psi}\boldsymbol{\delta}\right\|_1\right], \tag{2.50}$$

where $\boldsymbol{\Psi}$ is a sparsifying transform.

In contrast to Equation 2.49, the log-likelihood function in Equation 2.50 is in terms of the difference image, with a fixed contribution from the prior image. Note that for RoD reconstructions both penalties act on the difference image, whereas in PIRPLE, the sparsity term $\boldsymbol{\Psi}_2$ acts only on the follow-up

image. Also note that for RoD methods, care has to be taken to avoid non-negative reconstruction algorithms, since parameters can now be positive or negative [201].

2.4.3 Longitudinal Compressed Sensing MR Image Reconstruction

Prior images have also been used in MR image reconstruction problems. While more typically used for temporal regularisation in dynamic [204–206] and/or motion-resolved [207–210] MR, there has been research on prior image based reconstruction of longitudinal images of brain tumours [211, 212]. In these cases, the reconstruction objective function is similar to the PICCS (Equation 2.48) and PIRPLE (Equation 2.49) objective functions, trading off sparsity of the target image in some domain and sparsity of the difference image itself. However, MR imaging is more flexible than x-ray CT in that there is more control over the frequency sampling of images, since MR data is acquired directly in the Fourier domain. Therefore, these longitudinal MR methods also incorporate adaptive data sampling methods to optimise the undersampling pattern of the follow-up scans based on the baseline scan [211, 212].

2.4.4 Non-Negative Difference Image Reconstruction Methods in Emission Tomography

Similarly to the RoD method in CT, reconstruction of difference images in ET has also been a topic of research [213,214]. In this case though, the background image, $\boldsymbol{\theta}_B$, is estimated simultaneously to the difference image, $\boldsymbol{\delta}$, so that the model of the mean for the data ($\boldsymbol{q}_1$ and $\boldsymbol{q}_2$) is given by [213]:

$$\begin{bmatrix} \boldsymbol{q}_1 \\ \boldsymbol{q}_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{P}_1\boldsymbol{M}_{2\to 1} & \boldsymbol{0} \\ \alpha\boldsymbol{P}_2 & \alpha\boldsymbol{P}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta}_B \\ \boldsymbol{\delta} \end{bmatrix} \tag{2.51}$$

where $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are the system matrices for the two scans, α is a scaling factor to account for differing overall levels of detected counts, and $\boldsymbol{M}_{2\to1}$ is the operator that aligns the scan 2 image, $\boldsymbol{\theta}_2$, to the scan 1 image, $\boldsymbol{\theta}_1$. Note that in this formulation, the background image is assumed to be a transformed version of the scan 1 image such that $\boldsymbol{\theta}_1 = \boldsymbol{M}_{2\to1}\boldsymbol{\theta}_B$. Scan 2 is modelled as this transformed version of scan 1 plus the difference image $\boldsymbol{\delta}$. The Poisson log-likelihood objective function can then be formulated for the entire data in terms of $\boldsymbol{\theta}_B$ and $\boldsymbol{\delta}$.

However, using this model without any regularisation, it is clear that the ML solution of $\boldsymbol{\delta}$ must be the same as the difference of the ML solutions of $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_1$ separately, ignoring the interpolation effects of the alignment operator, and assuming that each scan has a unique ML solution. This is because once a suitable reconstruction algorithm has reached the unique ML solution of the background image, $\boldsymbol{\theta}_B^{\mathrm{ML}}$, which itself is just a transformed version of $\boldsymbol{\theta}_1$, the parameters in $\boldsymbol{\delta}$ may take any values in order to explain the $\boldsymbol{q}_2$ data vector. These values clearly will correspond to those that yield $\boldsymbol{\theta}_2^{\mathrm{ML}}$, i.e. the image solution had the two reconstructions been performed separately. Nonetheless, when using the non-linear EM algorithm to perform the joint reconstruction, as implemented in the past [213, 214], the inherent non-negativity constraint allows for apparent difference image improvement by disallowing negative values in the difference, thereby achieving visible noise reduction. This noise reduction aids detectability of regions of change [213, 214]. However, reduced lesion contrast in the difference images has been observed compared to independent MLEM reconstruction [213], probably due to a small positive bias in background regions with no change, in accordance with the low-counts bias observed with the standard MLEM reconstruction [132]. Furthermore, by formulating the model of the mean as in Equation 2.51, the convergence of the EM algorithm may be affected for the difference image, since there are correlations being introduced between parameters that must be resolved to approach the ML estimates.

2.4.5 Cross-Tracer SPECT Image Reconstruction

To perform a more meaningful simultaneous reconstruction, joint penalties can be defined to encourage the characteristics required from the resultant images. In this way, joint reconstruction of two simultaneously acquired radiotracers has previously been proposed in the context of myocardial perfusion single photon emission tomography (SPECT) imaging [215]. In this case, because the two images are acquired at the same time, misregistration was deemed to be negligible, allowing a so-called 'cross-tracer' prior term to be defined. The overall objective function was given by:

$$\Phi(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{m}, \boldsymbol{n}) = \Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m}) + \Phi^{\mathrm{ML}}(\boldsymbol{\phi}|\boldsymbol{n}) - \beta U(\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{2.52}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are the image vectors for the two radiotracers, with respective datasets denoted as $\boldsymbol{m}$ and $\boldsymbol{n}$, and $U(\boldsymbol{\theta}, \boldsymbol{\phi})$ is a joint penalty given by:

$$U(\boldsymbol{\theta}, \boldsymbol{\phi}) = \sum_{j=1}^{J} \sum_{k \in \mathcal{N}_j} w_{jk} \sqrt{1 + \left(\frac{\theta_j - \theta_k}{\delta}\right)^2 + \left(\frac{\phi_j - \phi_k}{\eta}\right)^2} - 1, \tag{2.53}$$

where δ and η are parameters to account for the varying intensity levels in the two scans. This prior is essentially a cross-tracer extension of the edge-preserving hyperbolic prior, designed to preserve edges that are present in both images.

2.4.6 Joint PET-MR Image Reconstruction

Simultaneous image reconstruction methods have also been applied to the multi-modality case, particularly for complementary PET and MR imaging. While the guided ET image reconstruction methods detailed above involve a one-directional transfer of information from the anatomical image to the ET image to be estimated, it may be the case that the ET image possesses relevant information not contained in the anatomical image, implying that a synergistic bi-directional transfer of information may be useful. With the advent of hybrid PET-MR scanners [216, 217] and the consistent aim in MR imaging to obtain

high-quality images with ever more highly undersampled (i.e. faster) data acquisition methods [218, 219], this synergistic use of information has been realised in joint PET-MR image reconstruction methods [148, 220–224]. In joint PET-MR image reconstruction, the objective function can be written as:

$$\Phi(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{m}, \boldsymbol{k}) = \Phi^{\text{PET}}(\boldsymbol{\theta}|\boldsymbol{m}) + \Phi^{\text{MR}}(\boldsymbol{\phi}|\boldsymbol{k}) - \beta U(\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{2.54}$$

where Φ^{PET} is the PET data fidelity term, i.e. the Poisson log-likelihood of the image $\boldsymbol{\theta}$ given the data $\boldsymbol{m}$, and Φ^{MR} is the MR data fidelity in terms of the image $\boldsymbol{\phi}$ and the MR data, $\boldsymbol{k}$ (usually a Gaussian noise model is used, leading to a least-squares fidelity term). $U(\boldsymbol{\theta}, \boldsymbol{\phi})$ is a penalty in terms of both of the images, that can be used to ensure a bi-directional transfer of information during the reconstruction process. This is usually to reduce noise and increase resolution in PET images, while reducing undersampling artefacts in MR images, allowing faster MR data acquisitions. Contrary to the anatomically guided ET image reconstruction literature, where there have been many different approaches proposed, nearly all of the joint PET-MR image reconstruction methods have used image spatial based priors to encourage common edges between the two modalities [148, 220, 221, 223, 224]. One exception to this is the work of Mehranian *et al* [222], which used a Bowsher-like weighted quadratic penalty method, with weights updated during the reconstruction based on the current image estimate.

Chapter 3: Simultaneous Reconstruction of Longitudinal PET by Encouraging Sparse Difference Images

Inspired by related work in both prior image constrained and joint image reconstruction, we define here a joint image reconstruction method for longitudinal oncology treatment response PET datasets that encourages sparsity in the difference image itself. The choice of sparsity-encouraging priors is motivated by similar observations made in previous work [199, 200], namely that in longitudinal oncology scans there is often only a small region of valid change, while a large part of the rest of the images remains similar (Figure 1.1). It is also important to note that since follow-up PET scans are often performed relatively early in order to image early functional changes, the level of expected change is lower than that observed in longitudinal MR, where previous research [211, 212] has had to use adaptive data acquisition to account for possible extreme changes. Furthermore, whereas diagnostic CT and MR generally allow the acquisition of low-noise prior images, PET images are often considerably noisy due to the limitations of PET data acquisition. This observation makes the prior-image constrained methods [199, 200, 211, 212] less suitable for PET imaging since such methods would be expected to transfer noise as well as genuine image information. For this reason we propose a simultaneous reconstruction approach that allows the improvement of both of a pair of longitudinal images.

In this Chapter, the concept of simultaneous reconstruction of longitudinal PET images is introduced, along with two priors to encourage sparsity in the difference image domain. These priors were assessed using 2D simulation studies and a 3D split-data inserted tumour study. This Chapter is based on work published by the author during the course of their PhD studies [225].

3.1 Sparsity-Encouraging Priors for Joint Longitudinal PET Image Reconstruction

For the reasons outlined above, a joint objective function with a sparsity-encouraging prior is defined for the simultaneous reconstruction of longitudinal PET datasets. To achieve this, the following joint PML objective function is defined in terms of two longitudinal image vectors, $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$:

$$\Phi(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{m}, \boldsymbol{n}) = \Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m}) + \Phi^{\mathrm{ML}}(\boldsymbol{\phi}|\boldsymbol{n}) - \beta U(\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{3.1}$$

where $\boldsymbol{n}$ denotes the measured data corresponding to the follow-up image vector $\boldsymbol{\phi}$.

In order to encourage sparsity in image reconstruction, a manner of quantifying the sparsity of a given signal must first be defined. Strictly speaking, sparsity is the number of non-zero elements that exist within a signal [226]. This concept can be formulated mathematically as the ℓ_0-"norm" (so-called because it does not satisfy the requirements of a true norm), given by:

$$\ell_0(\boldsymbol{x}) = \lim_{p \to 0} \sum_{j=1}^{J} |x_j|^p. \tag{3.2}$$

Despite allowing a mathematical definition of sparsity, the ℓ_0-"norm" is generally not suitable for use in PET image reconstruction problems for two reasons: firstly, the gradient of the ℓ_0-"norm" has a discontinuity at $x_j = 0$, meaning that it is not suitable for use in image reconstruction methods that rely on gradients. Secondly, the ℓ_0-"norm" is highly sensitive to noise, since

any small perturbation of a vector element from 0 results in that element being included in the count of non-zero elements.

For these reasons, two alternative forms for $U(\boldsymbol{\theta}, \boldsymbol{\phi})$ are proposed in this work in order to encourage sparsity in the difference domain (difference sparsity, DS); one convex and one non-convex (NC):

$$U_{\text{DS}}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \sum_{j=1}^{J} \sqrt{(\phi_j - \theta_j)^2 + \varepsilon^2} \tag{3.3a}$$

$$U_{\text{DS-NC}}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \sigma \sum_{j=1}^{J} \left[1 - \exp\left(-\frac{(\phi_j - \theta_j)^2}{\sigma^2} \right) \right]. \tag{3.3b}$$

Figure 3.1 shows the contribution from a single difference image voxel to these priors as a function of the value of the difference. The DS penalty in Equation 3.3a is a convex smoothed ℓ_1-norm penalty term to encourage sparsity, similar to the ℓ_1-norm priors justified theoretically for least-squares estimation problems by Donoho [195] and used widely in the image reconstruction literature (see Section 2.4 for examples). The second penalty, given by Equation 3.3b, is a non-convex function designed to more closely approximate the ℓ_0-"norm", while being tolerant to non-zero values arising from noise. In the DS-NC penalty the parameter σ determines the width of the Gaussian function relative to the ranges of differences observed and is also used as a scaling factor to keep the maximum magnitude of the one-step-late derivative independent of σ. Note that as σ tends towards zero, $U_{\text{DS-NC}}(\boldsymbol{\theta}, \boldsymbol{\phi})$ tends towards the ℓ_0-"norm". Conversely, with a non-zero σ, $U_{\text{DS-NC}}(\boldsymbol{\theta}, \boldsymbol{\phi})$ is convex only for a range of differences given by $-\sigma/\sqrt{2} \leq \phi_j - \theta_j \leq \sigma/\sqrt{2}$, and is approximately quadratic close to differences of zero.

In this Chapter the proposed longitudinal PML problem defined in Equation 3.1 is tackled using the OSL-MAPEM algorithm of Equation 2.33 [143]. This leads to the following update equations for $\boldsymbol{\theta}^{(\nu+1)}$ and $\boldsymbol{\phi}^{(\nu+1)}$:

$$\theta_j^{(\nu+1)} = \frac{\theta_j^{(\nu)}}{\sum_{i=1}^{I} P_{ij} + \beta \left. \frac{\partial U}{\partial \theta_j} \right|_{\boldsymbol{\theta}^{(\nu)}, \boldsymbol{\phi}^{(\nu)}}} \sum_{i=1}^{I} P_{ij} \frac{m_i}{\sum_{l=1}^{J} P_{il} \theta_l^{(\nu)} + b_i} \tag{3.4a}$$

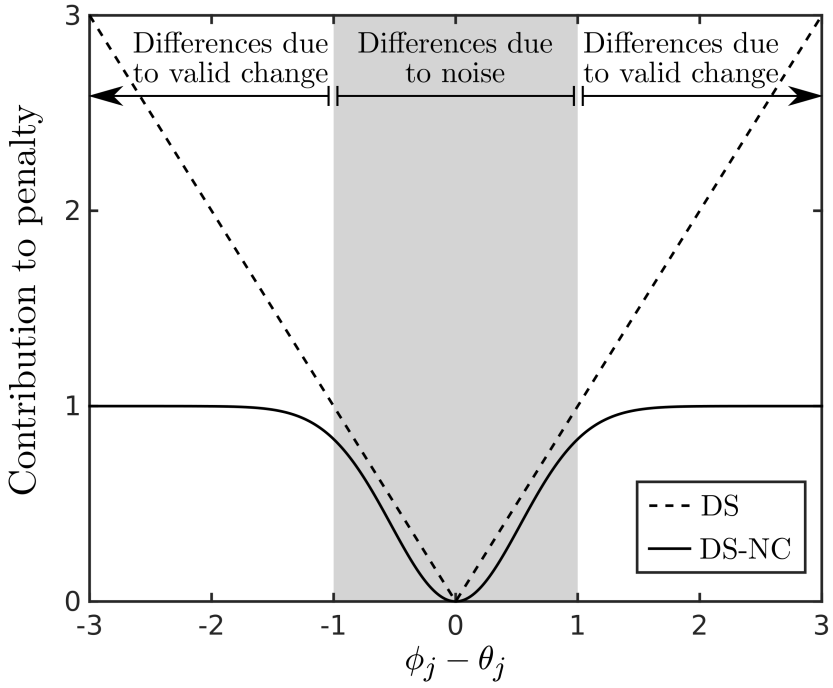


Figure 3.1: The contribution to the DS and DS-NC priors as a function of voxel-wise longitudinal change. Whereas the DS prior has a high gradient for all differences, including large ones, the gradient of the DS-NC prior tends to zero at large differences where the differences can be assumed to be due to valid change rather than noise-induced. Note that $\varepsilon = 0$ for the DS prior shown here.

$$\phi_j^{(\nu+1)} = \frac{\phi_j^{(\nu)}}{\sum_{i=1}^{I} Q_{ij} + \beta \left.\frac{\partial U}{\partial \phi_j}\right|_{\boldsymbol{\theta}^{(\nu)}, \boldsymbol{\phi}^{(\nu)}}} \sum_{i=1}^{I} Q_{ij} \frac{n_i}{\sum_{l=1}^{J} Q_{il}\phi_l^{(\nu)} + c_i}. \tag{3.4b}$$

where $\boldsymbol{Q}$ is the system matrix for the second PET dataset, $\boldsymbol{\theta}^{(\nu)}$ and $\boldsymbol{\phi}^{(\nu)}$ denote the image estimates after ν iterations and $\boldsymbol{c}$ is the estimated background counts vector for the second dataset.

When used with the DS prior (Equation 3.3a), the resulting reconstruction method is referred to as DS-PML, and when used with the DS-NC prior (Equation 3.3b), it is referred to as DS-NC-PML.

It should be noted that the methods presented in this section consider only the simplifying case where a) there are only two longitudinal datasets to be reconstructed, b) the two images have equal intensity in regions that have experienced no change in functional behaviour in the interval between scans, and c) the subject is in an identical position in both images so that there

is no misalignment between the two images. These simplifications are made throughout this Chapter for clarity as well as to provide the limiting case where the method would be expected to perform at its optimum. Extension of the theory to multiple scans and incorporation of these other effects is presented in Chapter 4.

3.2 2D Simulation Study

An initial 2D simulation study was performed in order to test the proposed simultaneous reconstruction methods with the two previously described priors. A static [^{18}F]-FDG PET phantom was used to simulate two PET acquisitions representing baseline (PET1) and follow-up (PET2) scans. These two datasets were reconstructed with the proposed DS-PML and DS-NC-PML methods and MLEM, and the resultant images evaluated with objective image quality metrics. Simulation of data, reconstruction of simulated data, and analysis of reconstructed images were carried out with MATLAB R2017a (MathWorks, MA, USA).

3.2.1 Experimental Methods

Data Simulation The freely available NCAT brain phantom [227] was used to create discretised 2D [^{18}F]-FDG activity and attenuation maps with a grid size of 512×512 pixels and an isotropic pixel size of side length 0.5 mm. By adding a number of hot spots to the initial activity map, the PET1 and PET2 ground truth images were defined (Figure 3.2(a-b)). This resulted in three regions of change in activity distribution between the scans: a frontal-right region where an active tumour embedded in white matter increases in size and activity, a mid-right region where an active tumour embedded in grey matter decreases in size and activity, and a frontal-left region where a tumour appears. It should be noted that since the DS-PML and DS-NC-PML methods do not distinguish the order of the two scans, PET1 and PET2 can be swapped without affecting the outcomes of the experiment.

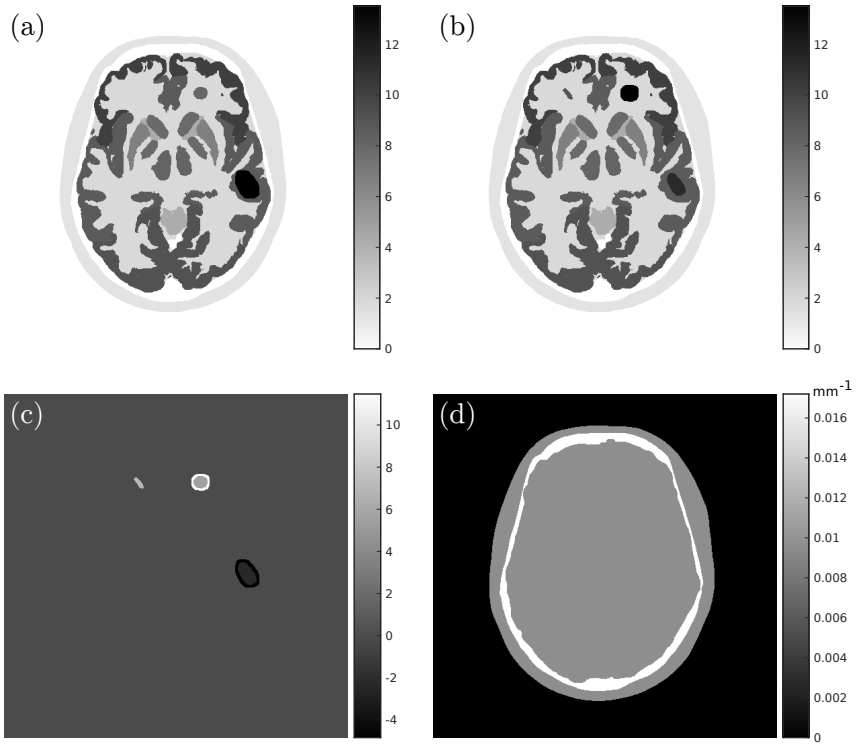


Figure 3.2: Ground truth images (a-c) and (d) the attenuation map used in the 2D simulation study. The ground truth images consist of (a) the PET1 activity map, (b), the PET2 activity map, and (c) the corresponding difference image displaying intensity changes due to changes in tumour sizes and activities.

PET data acquisition was simulated as follows: Firstly, resolution degradation was modelled as a single image-space Gaussian blur with a full width at half-maximum (FWHM) of 4.3 mm. These resolution-degraded images were transformed into sinogram space by forward projection using the Radon transform. Attenuation factors for 511 keV photons, μ, for each scan were defined using the attenuation map shown in Figure 3.2(d), with three pixel classes: air ($\mu = 0\,\mathrm{cm}^{-1}$), water ($\mu = 0.096\,\mathrm{cm}^{-1}$), and bone ($\mu = 0.172\,\mathrm{cm}^{-1}$) [106]. Randoms were simulated as uniform sinograms and scattered events were modelled as smoothed copies of the forward projection of each image. The number of expected counts in each of the noise-free sinograms was $\sim$2.2 M, with 20% from random and 20% from scattered events.

Reconstructions and Parameter Selection Simulated noisy PET sinograms were reconstructed into images using the attenuation factors and randoms and scatters sinograms used in the simulations. All images were reconstructed

into a 128×128 grid with a voxel size of $2{\times}2\,\mathrm{mm}^2$ and a uniform initial image estimate. Resolution modelling was not included in the reconstructions to avoid ringing artefacts [129]. For comparison purposes, the data were reconstructed using conventional MLEM [149] as well as the proposed DS-PML and DS-NC-PML methods. In addition, double-counts PET1 and PET2 datasets with an expected number of counts equal to $\sim$4.4 M were reconstructed with MLEM to serve as a reference.

For the DS-PML method, β values between 5×10^{-5} and 5×10^{-4} were used, with $\varepsilon = 1 \times 10^{-6}$. For DS-NC-PML, β values between 2.5×10^{-4} and 2.5×10^{-3} were used, with σ values ranging from 0.5 to 3.5. All reconstructions were run for 200 iterations.

Image Evaluation For a set of N Poisson noisy realisations of a simulated dataset, the reconstruction error relative to a reference image $\boldsymbol{\theta}^{\mathrm{Ref}}$ in a region Ω can be defined as a percentage root mean square error (%RMSE), given by:

$$\%\mathrm{RMSE} = \frac{100\%}{N_\Omega} \sum_{j \in \Omega} \frac{\sqrt{\frac{1}{N} \sum_{n=1}^{N} \left(\theta_{jn} - \theta_j^{\mathrm{Ref}}\right)^2}}{\theta_j^{\mathrm{Ref}}}, \tag{3.5}$$

where θ_{jn} is the jth voxel of the nth noisy realisation, and N_Ω is the number of voxels in Ω.

In this Chapter, noise-free FBP reconstructions of non-resolution-degraded ideal datasets were used as the reference images for all 2D simulation studies. This was done for the following reasons: Firstly, the use of noise-free FBP reference images guarantees that the reference images are in the same space as the reconstructed images, using an identical interpolation scheme in the projectors. For the MATLAB projectors used in this thesis this was seen to be very important, as attempts to produce the reference images simply by downsampling the ground truth images resulted in sub-voxel misalignment with the reconstructions. Secondly, using the noise-free FBP images means that the reference images include any projection errors, so that measuring %RMSE values of reconstructions relative to these references minimises the

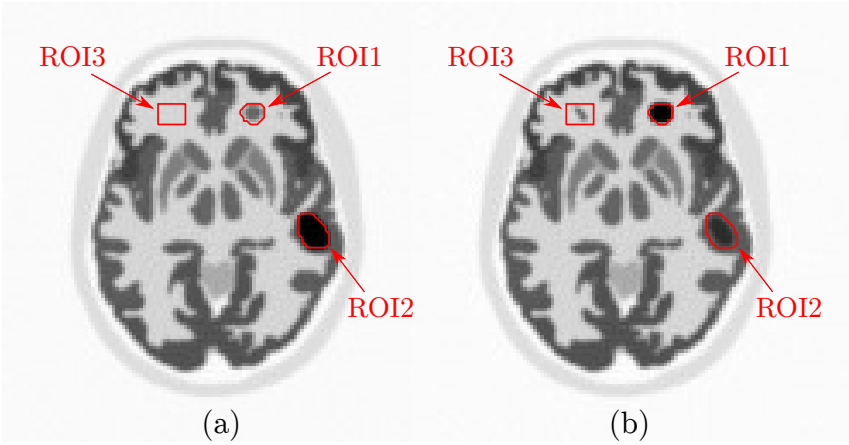


Figure 3.3: Reference images and tumour regions of interest (red outlines) for the 2D simulation study. (a) The PET1 reference image and (b) the PET2 reference image. These reference images were produced by performing FBP on noise-free, non-resolution-degraded data into the same image space as the noisy, resolution-degraded image reconstructions.

effect of these errors. Finally, the use of non-resolution-degraded data for the reference images allows sharp reference images, representing well the ground truth image. Sharper images that account for projector interpolation may be obtained by using noise-free MLEM reconstructions, but this would require the selection of the number of iterations used. Figure 3.3 shows the FBP reference images used for the 2D simulation study in this Chapter.

$N = 100$ noisy realisations of the data were used for reconstructions, and error levels were calculated across all non-zero voxels in the brain in the ground truth images (hereafter referred to as the whole-brain), as well as in ROIs around each of the three regions of change. Specifically, these ROIs were: the front-right tumour (ROI1), the mid-right tumour (ROI2), and the front-left tumour (ROI3). These latter three regions are shown in Figure 3.3 superimposed on the PET1 and PET2 reference images. Note that the same analysis regions were used in both the PET1 and PET2 reconstructed images, therefore including background voxels in the image where each tumour was smallest.

3.2.2 Results

Figure 3.4 shows the PET1 and PET2 %RMSE values in the four analysis regions for MLEM, double-counts MLEM, and DS-PML with various values of

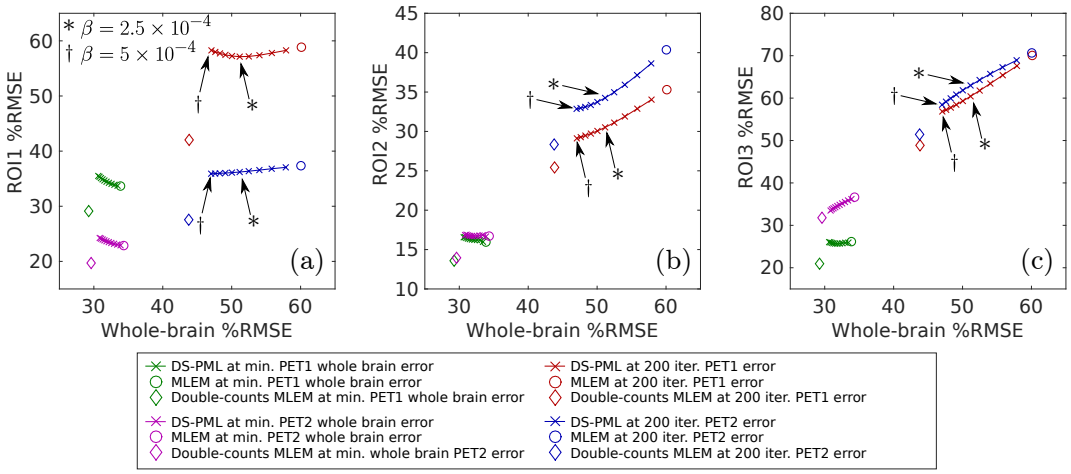


Figure 3.4: %RMSE values in (a) ROI1, (b) ROI2 and (c) ROI3 vs the whole-brain %RMSE for the 2D simulation study (see Figure 3.3 for the definition of each ROI). Circular markers show MLEM errors, diamond markers show error levels from MLEM with double the number of counts, and solid lines with cross markers show DS-PML errors as a function of β (β increases as cross markers move away from MLEM errors). For clarity, the DS-NC-PML results are displayed separately in Figure 3.5. Each graph contains PET1 errors at the minimum whole-brain error iteration (green), PET2 errors at the minimum whole-brain error iteration (magenta), PET1 errrors at iteration 200 (red), and PET2 errors at iteration 200 (blue). Arrows indicate error levels at 200 iterations for DS-PML with $\beta = 2.5 \times 10^{-4}$ ($*$) and $\beta = 5 \times 10^{-4}$ (†), corresponding to the DS-PML reconstructions shown in Figure 3.6.

β. Results are shown at iteration numbers corresponding to minimum whole-brain error and at 200 iterations. DS-PML reduces whole-brain error relative to MLEM both at minimum whole-brain error and at 200 iterations. As β increases, the error levels in the whole brain reduce, with the maximum β value of 5×10^{-4} reducing minimum whole-brain %RMSE by an average of 10% over both scans relative to MLEM. At 200 iterations, the corresponding improvement is 22%.

In the regions of change, the effect of the longitudinal regularisation is more varied. At the minimum whole-brain error iterations, %RMSE increases relative to MLEM with increasing β in ROI1 (Figure 3.4(a)) and PET1 ROI2 (Figure 3.4(b)). In ROI3 (Figure 3.4(c)), PET1 errors are stable with increasing β whereas the PET2 error reduces. However, at 200 iterations the error in the tumour regions generally decreases with DS-PML with increasing β, except for ROI1 in the PET1 reconstructions, where beyond $\beta = 2.5 \times 10^{-4}$ the error

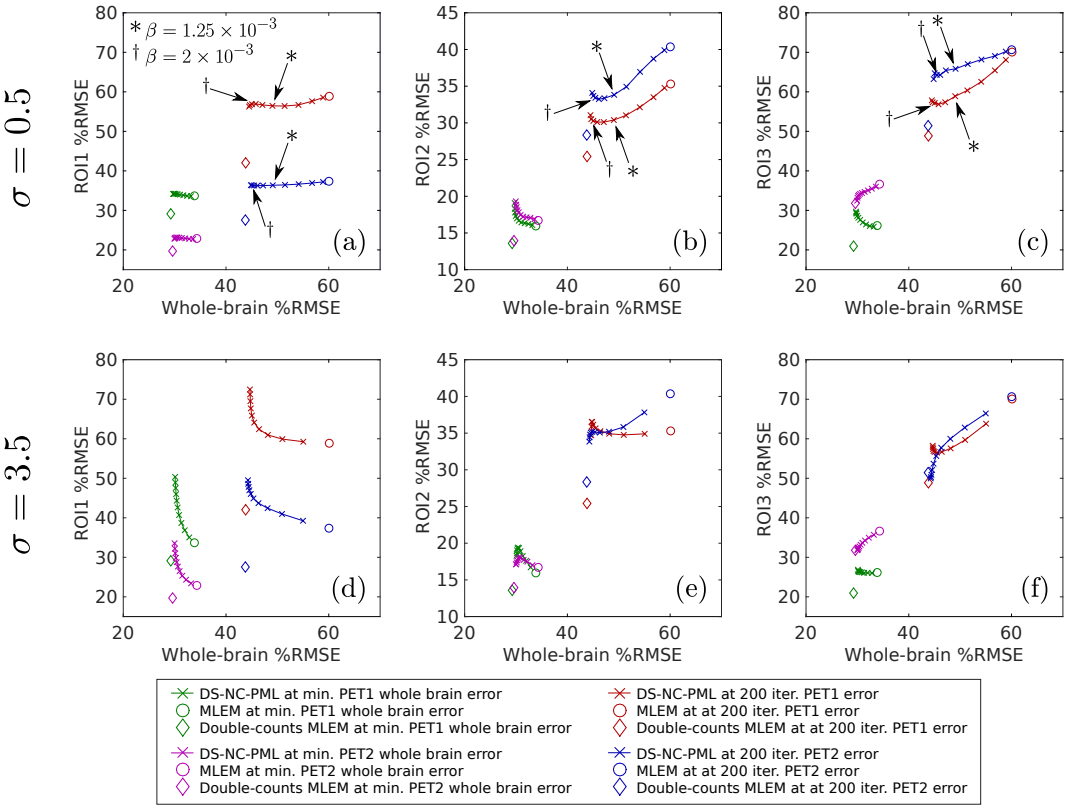


Figure 3.5: %RMSE tradeoffs (see Figure 3.4 for full description) for the DS-NC-PML method with $\sigma = 0.5$ (a-c) and 3.5 (d-f). Note that MLEM and double-counts MLEM results are replicated from Figure 3.4 for comparison purposes. Arrows indicate end iteration errors for DS-NC-PML with $\sigma = 0.5$ and $\beta = 1.25 \times 10^{-3}$ ($*$) and 2×10^{-3} (†), corresponding to the DS-NC-PML reconstructions displayed in Figure 3.6.

begins to rise slightly. In general, the increases in error levels in the regions of change are lower than the corresponding error reduction in the whole-brain error.

Figure 3.5 shows the error levels in the reconstructed images when using DS-NC-PML with σ values of 0.5 and 3.5. Similarly to the DS-PML case, the whole-brain error decreases with increasing β for both σ values. When $\sigma = 0.5$, %RMSE at minimum whole-brain error increases with β in ROI2 (Figure 3.5(b)) and in ROI3 in PET1 (Figure 3.5(c)), and decreases in the PET2 ROI3. In ROI1 (Figure 3.5(a)), the PET1 errors are reduced compared to the DS-PML results in Figure 3.4(a), both at minimum whole-brain error and at 200 iterations. Furthermore, although the general trends for the DS-NC-PML method are similar to the DS-PML results, the dependency on β is

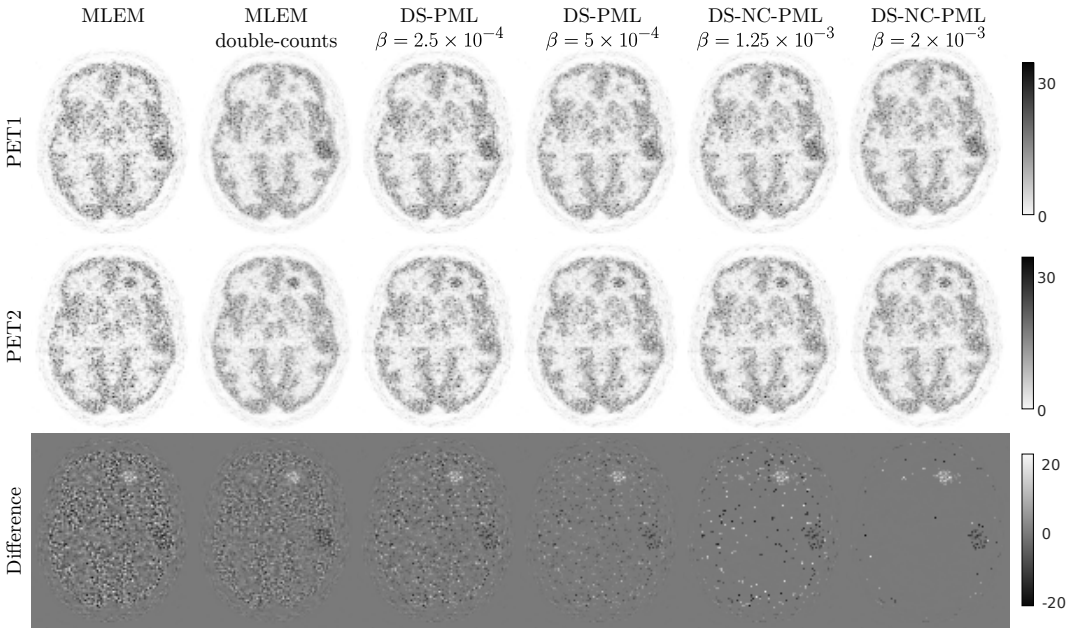


Figure 3.6: Example 2D simulation study single realisation reconstructed images after 200 iterations of MLEM, MLEM with double-counts, DS-PML with $\beta = 2.5 \times 10^{-4}$ and $\beta = 5 \times 10^{-4}$, and DS-NC-PML with $\sigma = 0.5$ and $\beta = 1.25 \times 10^{-3}$ and $\beta = 2 \times 10^{-3}$. Top row: PET1, middle row: PET2, bottom row: the difference PET2−PET1. When using the proposed simultaneous reconstruction methods, noise reduces relative to MLEM across the image, with noise levels visually approaching those observed when using MLEM with double count data. Furthermore, the proposed priors produce sparser difference images.

more non-linear. For β values beyond $\approx 1.25 \times 10^{-3}$, the behaviour changes, with errors at minimum whole-brain error changing more rapidly in ROI2 and ROI3. In fact, this change in behaviour causes the 200-iteration error to start to rise with higher β values for ROI2 in PET1 and PET2, and ROI3 in PET1. Overall, with β values up to $\approx 1.25 \times 10^{-3}$, the DS-NC-PML method with $\sigma = 0.5$ performs similarly to DS-PML, but with slight error reduction in ROI1.

Increasing σ to 3.5 degrades the performance of the DS-NC-PML method (Figure 3.5(d-f)). The ROI1 error (Figure 3.5(d)) increases with β in both PET1 and PET2 at minimum whole-brain error and at 200 iterations. In addition, the PET1 ROI2 error (Figure 3.5(e)) at 200 iterations increases relative to MLEM with any value of β, as well as errors at minimum whole-brain error rising faster with β than for $\sigma = 0.5$.

Example reconstructed images after 200 iterations of each method are shown in Figure 3.6 for MLEM, double-counts MLEM, DS-PML with $\beta = 2.5 \times 10^{-4}$ and $\beta = 5 \times 10^{-4}$, and DS-NC-PML with $\sigma = 0.5$ and $\beta = 1.25 \times 10^{-3}$ and 2×10^{-3}. The selected regularisation parameters avoid excessive penalisation of the regions of change whilst demonstrating the benefits of the respective methods. For the DS-NC-PML method, the two β values were used to demonstrate the more complex relationship between error levels and β observed in Figure 3.5. Figure 3.6 demonstrates that using the DS-PML and DS-NC-PML methods improves the reconstructed images by reducing noise throughout the brain, approaching noise levels observed with double-counts MLEM. This is particularly clear in and around the striatum and is reflected in the difference images, where the amplitude of the background noise reduces when using the proposed methods. Note that according to Figures 3.4(a) and 3.5(a) the displayed DS-NC-PML reconstruction with $\beta = 1.25 \times 10^{-3}$ achieves on average slightly lower whole-brain error than DS-PML with $\beta = 2.5 \times 10^{-4}$ (49.1 and 51.2 respectively) while also achieving lower PET1 error in ROI1 (56.5 and 57.1 respectively). In addition, the corresponding DS-NC-PML difference image is sparser than the DS-PML. Furthermore, increasing the β value to 2×10^{-3} in the DS-NC-PML method produces an even sparser difference image; however the error levels in ROI3 at this penalty strength begin to rise again (Figure 3.5(c)).

3.3 3D Real Data Inserted Tumour Study

Following the 2D simulation study, the proposed DS-PML method was applied to real 3D data from a single [^{18}F]-FDG scan of a suspected frontal lobe epilepsy patient acquired on a Biograph mMR PET-MR scanner [217] (Siemens Healthcare, Erlangen, Germany). The patient was injected with a total activity of 182.9 MBq and scanned for 30 min at 1.3 hr post-injection, resulting in a recorded total of ∼600 M prompt counts. Attenuation factors were estimated using an MR-based attenuation map generated from vendor-provided ultra-short echo time sequences. Randoms and scatters were estimated with

vendor-supplied tools using the delayed coincidence method and multi-slice 2D single scatter simulations respectively. Vendor-supplied normalisation files were also used in the reconstructions. Reconstructions and image analysis were carried out in MATLAB, with 3D mMR projections performed using in-house software described previously [228, 229].

3.3.1 Experimental Methods

Dataset Generation To create two pseudo-longitudinal datasets representing the PET1 and PET2 scanning sessions, the sinogram from the patient scan was split equally by randomly assigning each prompt count to one of two new datasets with equal probability. This resulted in two 3D Poisson-distributed emission datasets whose sum was equal to the original data. Attenuation and normalisation factors for each of the split datasets were the same as for the original full-counts dataset. Randoms and scatters estimates for the split-data sinograms were defined to be 50% of the original estimates.

A high-activity spherical tumour was then introduced to the same position in each dataset. For the PET1 dataset the tumour had a radius of 6.0 mm and additive intensity of 0.3 (equivalent to ∼120% the value of surrounding grey matter). For the PET2 datasets, three tumour radii of 4.5 mm, 6.0 mm and 7.5 mm were defined, with additive intensities of 0.15, 0.30, or 0.45. By using all combinations of intensity and size, nine PET2 tumours were simulated, allowing investigation of nine cases of longitudinal change.

Tumours were introduced into the pseudo-longitudinal datasets by forward projecting tumour-only images, applying attenuation and normalisation factors, and then introducing Poisson noise. Randoms and scatters for the tumours were not calculated or included in the generation pseudo-longitudinal datasets. The resulting Poisson noisy tumour-only sinograms were then added to corresponding split-data sinograms, resulting in Poisson noisy sinograms including tumours. Randoms and scatters estimates were not re-calculated using the final inserted tumour sinograms.

Reconstructions The pseudo-longitudinal sinograms for each case of tumour change were reconstructed with both MLEM and DS-PML with up to 100 iterations, initialised with uniform images. For the DS-PML method, β was set to a value of 20 (determined empirically, since the 2D β values are not directly applicable to the 3D case), and ε was set to 0. No resolution modelling was used in the reconstructions. In addition, double-counts MLEM images were obtained for each case of tumour change by inserting tumours into the original dataset and reconstructing with 100 iterations of MLEM. All images were reconstructed into a $344 \times 344 \times 127$ grid with a voxel size of $2.08626 \times 2.08626 \times 2.03125$ mm^3.

Image Evaluation Images were analysed by measuring noise levels in two approximately uniform regions of the reconstructed images. The first region was a 64 voxel region in the white matter and the second was a 32 voxel region within the occipital grey matter. Noise in a given uniform region Ω for a given image $\boldsymbol{\theta}$ was quantified by using the coefficient of variation (CV) according to

$$\mathrm{CV}_\Omega(\boldsymbol{\theta}) = \frac{\mathrm{SD}_\Omega(\boldsymbol{\theta})}{\mathrm{Mean}_\Omega(\boldsymbol{\theta})} \tag{3.6}$$

where $\mathrm{SD}_\Omega(\boldsymbol{\theta})$ is the sample standard deviation of voxel values of $\boldsymbol{\theta}$ in the region Ω, and $\mathrm{Mean}_\Omega(\boldsymbol{\theta})$ is the mean value in the same region.

In addition, the mean values in the tumour, white matter region, and grey matter region were recorded in PET1 and PET2 in order to observe the level of bias introduced by the DS-PML method. Note that reconstructed image intensities using double-counts MLEM were halved prior to analysis to provide comparable regional means.

Ten realisations of the random data-splitting process were performed for each tumour change and the mean CV and mean values across realisations were used as final figures of merit.

Counts Reduction Experiment In order to investigate the effect of the counts level on performance, a single case of tumour change from the previous experiment was selected and reconstructed with MLEM and DS-PML for various

levels of counts per dataset. Specifically, the tumour change was a radius reduction from 6.0 mm to 4.5 mm, with an additive intensity reduction from 0.30 to 0.15 arb. units (as described above). The original emission dataset was split into pairs of datasets as previously described, with varying levels of probability such that the pairs of output datasets contained between 1% and 50% of the original ∼600 M counts each. The appropriately scaled tumours were then inserted into each of the datasets, resulting in a number of datasets nominally of the same longitudinally changing object with varying recorded counts levels.

Each of these pairs of datasets was reconstructed with 100 iterations of both MLEM and DS-PML with $\beta = 20$ and $\varepsilon = 0$. In order to analyse the PET1 and PET2 image quality across counts levels, the contrast to noise ratio (CNR) between the tumour and a background region in the adjacent grey matter was defined as:

$$\mathrm{CNR}(\boldsymbol{\theta}) = \frac{\mathrm{Mean}_T(\boldsymbol{\theta}) - \mathrm{Mean}_B(\boldsymbol{\theta})}{\mathrm{SD}_B(\boldsymbol{\theta})} \tag{3.7}$$

where $\mathrm{Mean}_T(\boldsymbol{\theta})$ denotes the mean tumour intensity within the mask used to define the tumour, and $\mathrm{Mean}_B(\boldsymbol{\theta})$ and $\mathrm{SD}_B(\boldsymbol{\theta})$ denote the mean and standard deviation in the background region respectively. A higher CNR is assumed to indicate a superior image quality since it indirectly measures the extent to which the tumour is visible against the surrounding tissue. Five realisations of the data generation process were performed for each counts level and the average CNR was calculated across realisations.

3.3.2 Results

Figure 3.7 shows example grey and white matter regional CV vs mean values for a range of iteration numbers for the 3D PET1 reconstructions using MLEM, DS-PML, and double-counts MLEM, for the case where the tumour radius reduced to 4.5 mm and the additive intensity reduced to +0.15. The mean values in these regions are improved by the longitudinal penalty, with

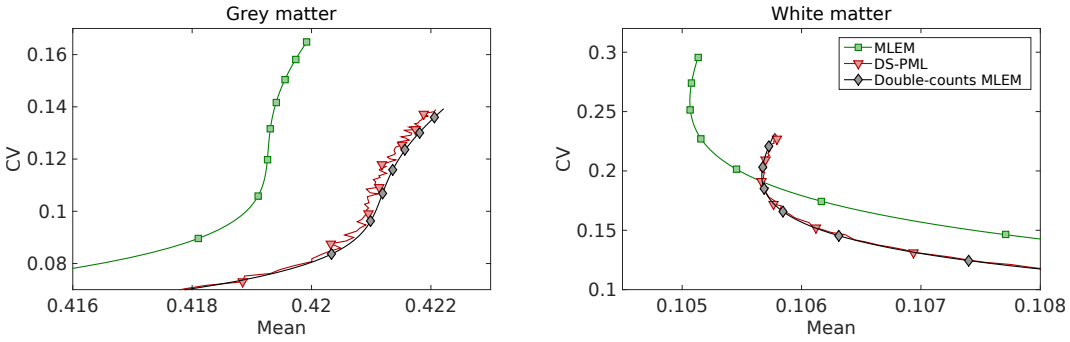


Figure 3.7: PET1 regional grey matter and white matter CV and mean as a function of iteration number for the 3D split data study for: MLEM (green squares), double-counts MLEM (grey diamonds) and DS-PML with $\beta = 20$ (red triangles). Note that these results are only for PET1 reconstructions for the case where the tumour radius reduced to 4.5 mm and the additive intensity reduced to +0.15; similar trends were observed for other reconstructions.

DS-PML producing regional means that agree with the double-counts MLEM reconstruction. In addition, using DS-PML reduces the noise compared to standard MLEM, with CV values very close to double-counts MLEM at all iteration numbers. Note that although the DS-PML curves in Figure 3.7 exhibit some oscillatory behaviour in tumour mean values arising from the OSL algorithm used, these oscillations are small compared to the absolute values under consideration.

These improvements are evident in the corresponding reconstructed images (Figure 3.8). Image-wide noise at 100 iterations of DS-PML is visibly reduced when compared to 100 iterations of MLEM with the same split data, to the point where the DS-PML reconstructions are visually similar to the double-counts MLEM images. In addition to demonstrating reduced image noise, Figure 3.8 shows that the visual appearance of the tumour is unaffected by the applied longitudinal regularisation. Inspection of the difference images confirms this observation, with the region of change remaining clearly visible.

In terms of the quantification of the tumour, Figure 3.9 shows the tumour means along with white matter CVs for all tumour changes for all methods at 100 iterations. For the case corresponding to Figure 3.8, where the tumour reduces in radius to 4.5 mm and additive intensity to +0.15, a slight bias is introduced into the DS-PML reconstructions so that the PET1 tumour mean

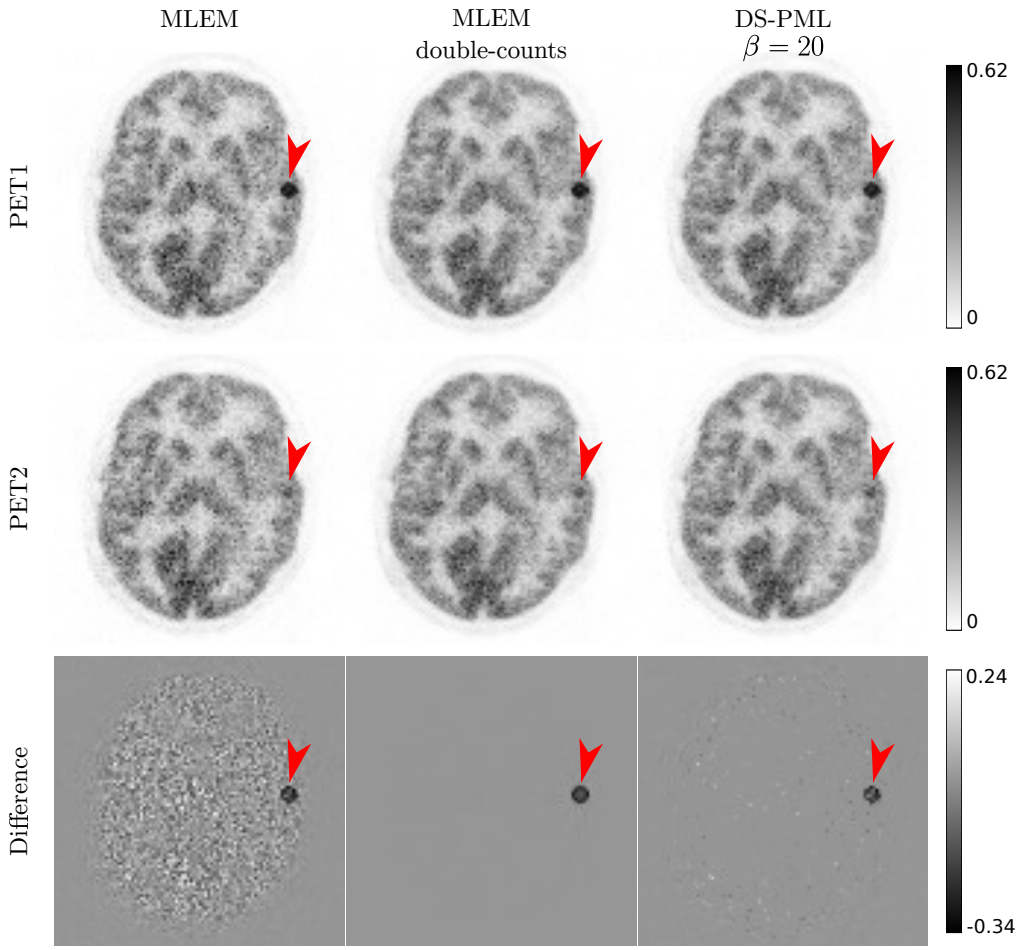


Figure 3.8: Reconstructed images at 100 iterations for the 3D tumour change experiment for the case where the tumour radius reduced to 4.5 mm and the additive intensity reduced to +0.15. Displayed reconstruction methods are MLEM, double-counts MLEM, and DS-PML with $\beta = 20$. Top row: PET1, middle row: PET2, bottom row: the difference PET2−PET1. Arrows indicate the location of the inserted tumour. Note that for the double-counts MLEM case the same noisy data constituted both the PET1 and PET2 datasets and so the difference image contains only the change due to the inserted tumour.

is underestimated by 2.2% and the PET2 tumour mean is overestimated by 2.6% relative to the double-counts MLEM reconstructions.

For the other tumours the regional means are either slightly biased towards the longitudinal average, or they agree with the double-counts MLEM value. The largest observed percentage bias relative to the double-counts reconstruction was −6% for the PET1 tumour when the PET2 tumour was 7.5 mm in radius and 0.15 in additive intensity, with all other bias magnitudes less than 5%.

Figure 3.10(a) shows the measured PET1 and PET2 tumour to background

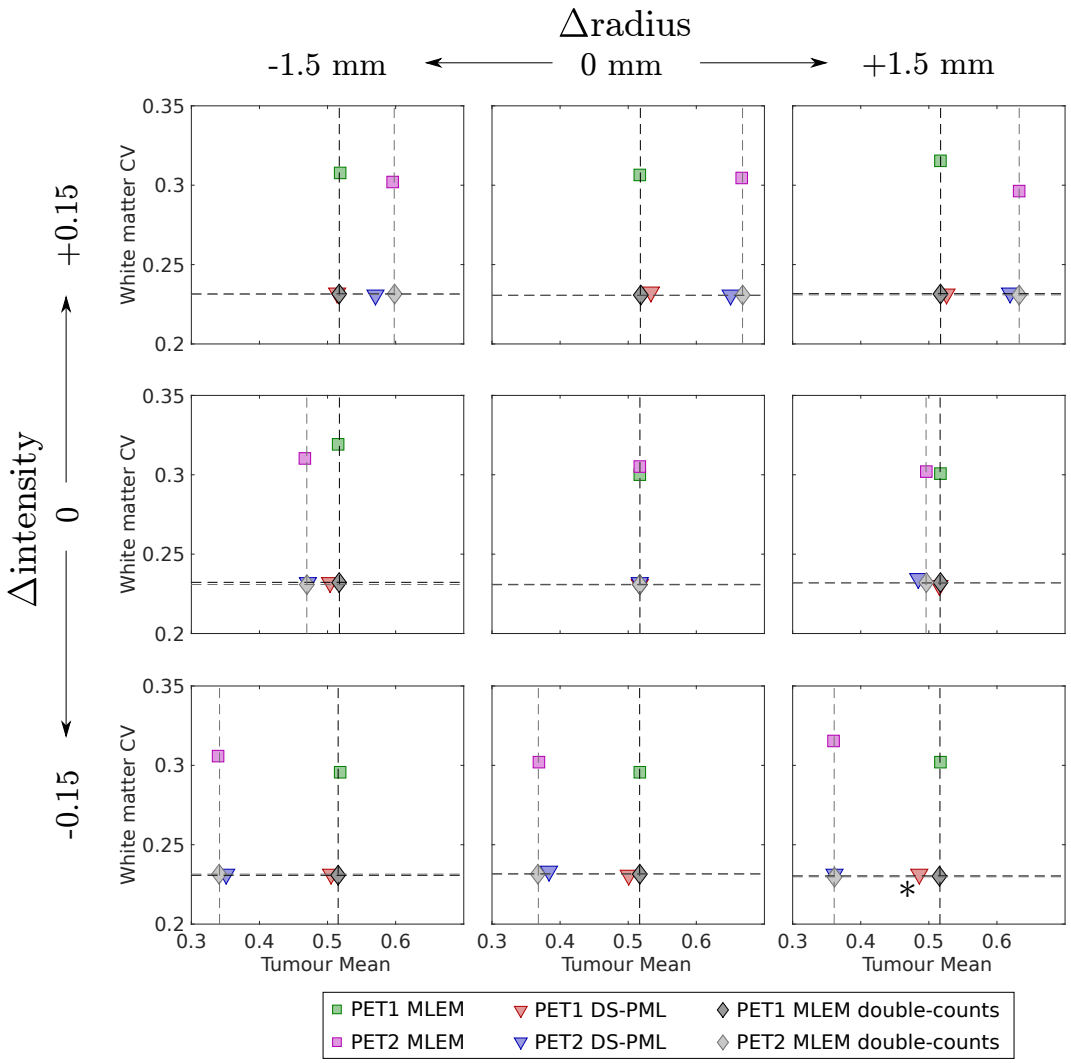
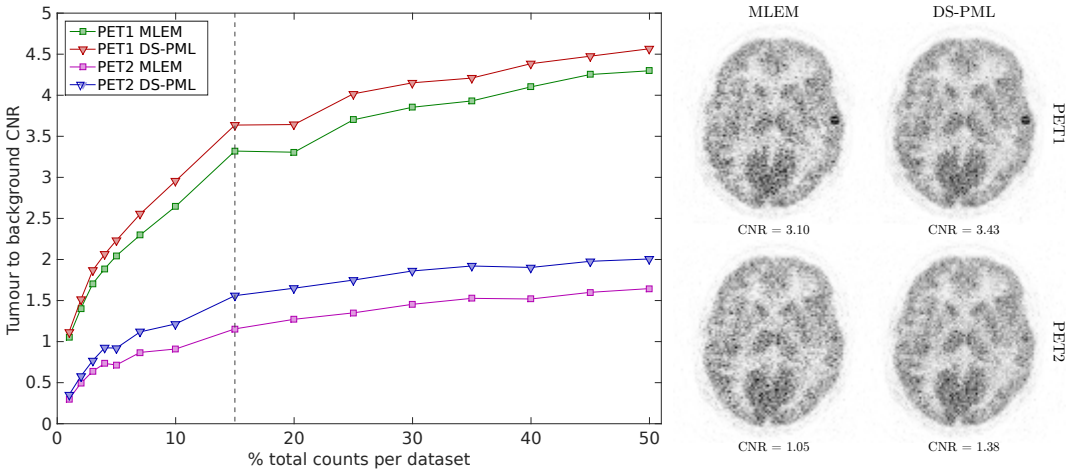


Figure 3.9: Tumour mean and white matter CV values for all nine tumour changes at 100 iterations. Each change is a combination of PET2 tumour additive value and radius, with each plot showing the results for PET1 and PET2 using MLEM (squares), MLEM with double-counts data (diamonds), and DS-PML with $\beta = 20$ (triangles). In all cases, the PET1 dataset included an inserted tumour of radius 6 mm and an additive value of +0.30. The largest observed tumour bias (relative to double-counts MLEM) using DS-PML was of -6% in the PET1 tumour when the PET2 tumour radius increased by 1.5 mm and the intensity decreased by 0.15 (bottom right, $*$). In all cases, the noise in the white matter region was reduced to the same level as the double-counts MLEM reconstruction when using DS-PML with 50% of the counts.



(a) (b)

Figure 3.10: Results of the counts reduction experiment showing (a) the measured CNR at various counts levels per dataset and (b) example reconstructed images at the 15% counts per dataset noise level (dashed line in (a)), using identical datasets for MLEM and DS-PML with $\beta = 20$. Labels in (b) show the measured CNR in each of the displayed images.

CNRs for count levels equal to 1% of the original dataset up to 50% (i.e. the case shown in Figure 3.8) for MLEM and DS-PML. At all noise levels, and for both PET1 and PET2 reconstructions, DS-PML was observed to produce a higher CNR than the corresponding MLEM reconstructions. Visual inspection of reconstructed images at the 15% counts level case (Figure 3.10(b)) reflects these improvements, with the tumour more visible in the DS-PML reconstructions due to a reduction of noise in both the adjacent background region and across the images in general.

3.4 Discussion

This Chapter has presented a new PET image reconstruction approach that permits simultaneous reconstruction of longitudinal PET images. By penalising voxel-wise differences between estimated images during the reconstruction, common information is transferred between the images.

The proposed methods were tested in a 2D simulation study which demonstrated that applying longitudinal regularisation improved the reconstructed images compared to MLEM by reducing whole-brain error while maintaining

or even reducing error levels in the regions of change. The decrease in whole-brain error is due to the noise reduction achieved by penalising longitudinal differences. By penalising voxel-wise differences between PET1 and PET2 images, information from both datasets was used in the reconstruction of both images, in effect increasing the number of counts available for each scan. Furthermore, by performing this penalty using sparsity-encouraging priors the output difference images were sparser than those produced using traditional MLEM reconstruction.

Importantly, it was observed that the error levels in the two right tumours (ROI1 and ROI2) at minimum whole-brain error rose with increasing β using the DS-PML method, essentially creating a tradeoff between error in the whole-brain and error in regions of change. This increased error in regions of change is because the DS prior penalises all differences with the same force (gradient of the prior term), resulting in some degradation of the longitudinal differences caused by tumour change. On the other hand, ROI3 in the PET2 reconstruction benefited from the DS prior due to improved reconstruction of the surrounding white matter.

Another observation of interest is that the 200-iteration DS-PML reconstruction error reduced in the mid-right tumour (ROI2) but increased slightly in the front-right PET1 tumour (ROI1) at high β values. This may be due to the specific bias-noise behaviour of each tumour, such that in the mid-right tumour where the magnitude of the ground truth change is 2.25, (i.e. relatively low), the noise reduction is stronger than the bias, whereas for the front-right tumour where the corresponding change magnitude is 5.4, the bias-induced error overrides the noise reduction.

When using the DS-NC-PML method with $\sigma = 0.5$, whole-brain error also decreased due to the effective sharing of information between scans, similarly to when using the DS-PML method. However, the error levels at minimum whole-brain error were not as adversely affected in regions of change as for DS-PML in ROI1. The reason for this is that the non-convexity of the prior provides a derivative of zero (i.e. zero force) at high changes. In terms of the iterative update in Equations 3.4a and 3.4b, this means that voxels with

a large longitudinal change revert to MLEM updates, and the information transfer between the scans in these voxels is reduced.

However, as shown in Figure 3.5, the DS-NC-PML method is sensitive to the values of β and σ. Reconstruction errors can be more non-linear in β than they are for DS-PML (Figure 3.5(b-c)), and with an excessive σ value the reconstruction penalises valid differences more and the error levels rise substantially (Figure 3.5(d-e)). Determining an *a priori* good value of σ is not necessarily obvious though, since it will depend on image noise levels and the magnitude of changes that are to be reconstructed.

In addition, the non-convexity of the DS-NC-PML makes it a theoretically less robust option for reconstructing longitudinal datasets because the use of a non-convex prior can result in a number of local maxima in the objective function. The number and nature of these local maxima would depend on the sizes of σ and β, and it would be difficult to ensure that globally optimum parameter estimates were being approached for the reconstruction of any given longitudinal data series.

In the results for the 3D data experiment, using DS-PML reduced image-wide noise to similar levels achieved when the number of counts was doubled and reconstructed with standard MLEM. In general, for MLEM reconstructions, noise reduces when the counts increase due to the signal-to-noise properties of Poisson distributions. The fact that DS-PML reduced background noise to similar levels as doubling counts suggests that the method is in effect sharing counts between datasets.

The amount of bias introduced in the tumour varied depending on tumour size and intensity, with the largest observed bias of -6% in the case of increasing tumour size and decreasing tumour activity. In practice, the performance at a given tumour change is likely to depend on the strength of the regularisation, making future validation of these effects over a greater range of realistic changes important.

As the number of counts per dataset reduced from 50%, DS-PML continued to provide improved images, measured as an increase in tumour to background CNRs compared to MLEM. There is the implicit assumption that the CNR

accurately reflects the overall image quality, which may not be true when biases are present. For example, the results of the tumour change experiment (Figure 3.9) show that longitudinal bias is possible when using the DS-PML method. For the PET2 reconstructions, this bias was generally observed to be positive, with tumour means increasing relative to the MLEM values. Using Equation 3.7, it is clear that an increase in CNR can occur with either an increase in tumour mean or a decrease in background noise (given that the background mean is stable). Therefore, for PET2 reconstructions, the increased CNR could be due to a combination of both background noise reduction and the longitudinal bias in the tumour. However, the PET1 CNR values are also increased by using DS-PML compared to MLEM. In these images, tumour means should only be biased negatively, lowering the CNR. The observation that PET1 CNR rises even with the presence of this negative longitudinal bias suggests that the longitudinal biases are small relative to the noise reduction, and also suggests that the CNR improvements seen for PET2 are primarily due to the noise reduction in the background as desired.

In the experiments performed in this Chapter, the proposed methods were compared to MLEM reconstructions of the same data, and also MLEM reconstructions of double-counts data to provide a high-counts reference reconstruction. There is, however, another simple potential comparison method that would be useful to include: the naïve longitudinal smooth. In the two-scan case, this would correspond to a weighted voxel-wise average of the two images. Since this method would also provide noise reduction levels between MLEM and double-counts MLEM reconstructions (depending on the weights used), it would better contextualise the biases reported in this Chapter when using DS-PML and DS-NC-PML. For these reasons, the longitudinal smooth is introduced as a comparison method in Chapter 4.

Another weakness of the experimental results provided in this Chapter is the reliance on %RMSE alone as the image quality metric in the presence of a known ground truth image in the 2D simulation study. It would be informative to investigate the bias and variance components of the overall error separately, to fully characterise the tradeoff between the two in regions

of change. This issue is addressed in Chapter 4, where additional simulation studies are performed and bias-variance analysis is performed.

An additional point that was not addressed in this Chapter was the role of resolution modelling. Including resolution effects in the forward model in the reconstruction, as in Equation 2.10, generally reduces image noise compared to non-resolution-modelled reconstructions (at the potential cost of ringing artefacts). Therefore, the reconstructed images displayed in this Chapter (both for 2D simulation studies an 3D real data inserted lesion study) could be further denoised by including resolution-modelling in the reconstruction. Nonetheless, while this would reduce the noise (CV) in each reconstructed image, the overall results of this Chapter would still be expected to hold, namely that the use of the proposed penalties reduces noise *relative* to the standard independent-dataset reconstructions. It may be expected that the relative improvement of the proposed methods would be lower in this case, although this would need to be checked in future work.

Despite these issues, the results of both the 2D simulated and 3D split real data studies presented in this Chapter do show that there is generally a trade-off between reconstruction quantification performance in regions of change and regions that do not change when using the proposed simultaneous reconstruction methods. The most suitable level of compromise will in practice be largely application-specific; metastasis localisation studies may be more tolerant to biases induced in lesions and benefit more from the increased CNR of the proposed methods, while studies where accurate quantification is of interest may benefit only a small amount (if at all) from simultaneous reconstruction. As such, it is clear that the proposed methods will require application-specific testing before being used in a clinical setting to ascertain a) if there is a benefit to using the proposed reconstruction methods in those contexts, and b) which level of regularistion is optimal.

From a technical point of view, the methods proposed in this Chapter have some limitations. Firstly, the DS-PML and DS-NC-PML methods as presented here used the assumption that the majority of voxels are equal in intensity in both the PET1 and PET2 images. In practice, the injected activity of

radiotracer in the PET2 acquisition is not necessarily equal to that of the PET1 acquisition. In this case, correcting each image for the relative injected activity would be necessary to produce difference images that are predominantly zero. This limitation is addressed in Chapter 4

Another simplification made in the work presented in this Chapter was the absence of misalignment between longitudinal images. In real data applications there is always some level of misalignment which would require correction to use DS-PML and DS-NC-PML. This could be overcome either through using side information such as CT or MR images to estimate the alignment operator, or by estimating the alignment as part of the simultaneous longitudinal image reconstruction, similarly to that used previously in both PET [230–233] and CT [199, 200] image reconstruction. In Chapter 4, alignment operators calculated from side information are introduced to overcome this limitation.

One final limitation of the methods proposed in this Chapter is their restriction to pairs of scans only. In reality, longitudinal treatment response PET may entail many more scans [66, 67, 73, 74], which would necessitate extension of the methods proposed in this work. Including even more scans in the reconstruction would be expected to further improve the performance of the proposed methods by making available a greater amount of information to be included in the reconstruction of each image. This idea is also developed in Chapter 4.

Nevertheless, the results presented in this Chapter demonstrate that it is in principle possible to improve reconstructed PET images by penalising voxel-wise longitudinal differences.

3.5 Summary

In this Chapter we proposed a novel penalised PET image reconstruction method to reconstruct longitudinal oncology treatment response PET images simultaneously. The large degree of similarity between these datasets allows for information to be exchanged longitudinally in order to improve the reconstructed images via noise reduction. This was achieved by penalising differ-

ences between two longitudinal images during reconstruction in a one-step-late fashion, using both ℓ_1-norm and non-convex difference penalties. The proposed methods, DS-PML and DS-NC-PML, were evaluated on 2D simulated datasets where lower whole-brain error levels were observed compared to MLEM, with error levels similar in longitudinally changing tumours. In real 3D split datasets with inserted tumours, an appropriate choice of hyperparameters for DS-PML reduced noise to levels similar to those normally observed by using twice the number of counts, while also maintaining quantification of tumour means.

Limitations of the proposed methods included the lack of global intensity corrections to account for varying injected activity levels, the lack of the misalignment step crucial for use on real data, and the restriction to two scans. These issues are addressed in the Chapter 4, along with the proposal of alternative difference-image based penalties designed to exploit other common characteristics of longitudinal difference images.

Chapter 4: Generalisation to Multiple Scans and Alternative Difference-Image Penalties

The results presented in Chapter 3 demonstrated that coupling a pair of longitudinal treatment response PET datasets in the reconstruction process allowed for noise reduction in both scans, reduced reconstruction error, and increased tumour-to-background CNR. However, some bias was introduced into the observed value of the tumour mean uptake, and only the specific case of two perfectly aligned scans with equal injected PET activity was considered.

In this Chapter, we extend the theory presented in Chapter 3 to a general number of scans with the necessary alignment and dose-correction steps, and present two more difference-image penalties designed to achieve similar levels of noise reduction with reduced levels of tumour bias. This Chapter is based on work published by the author during the course of their PhD studies [234].

4.1 Generalisation to S Scans and Accounting for Misalignment and Variable Activity Levels

Following on from the reconstruction problem defined in Section 3.1, here the difference-image based simultaneous longitudinal PET image reconstruction framework is generalised in order to a) encompass an arbitrary number of

scans, b) account for misalignment between the scans, c) account for changing injected activity, and d) allow a wider range of difference-image penalties to be easily incorporated into the simultaneous reconstruction framework.

As defined previously, $\boldsymbol{\theta}$ is a vector containing the intensities of a PET image so that θ_j is proportional to the activity concentration in the jth voxel. In this Chapter, in order to consider a general number of scans, a set of S longitudinal PET images is denoted as $\{\boldsymbol{\theta}_s\}$, where $s = 1, ..., S$ indicates the scan number. In this notation, the intensity of the jth voxel of the sth longitudinal scan is denoted as $\theta_{j,s}$.

Firstly, the joint objective function given in Equation 3.1 is generalised to include S scans by re-writing it as:

$$\Phi(\{\boldsymbol{\theta}_s\} \mid \{\boldsymbol{m}_s\}) = \sum_{s=1}^{S} \Phi^{\mathrm{ML}}(\boldsymbol{\theta}_s|\boldsymbol{m}_s) - \beta U(\{\boldsymbol{\theta}_s\}) \tag{4.1}$$

where $\Phi^{\mathrm{ML}}(\boldsymbol{\theta}_s|\boldsymbol{m}_s)$ is the Poisson log-likelihood for image $\boldsymbol{\theta}_s$, $U(\{\boldsymbol{\theta}_s\})$ is a penalty function in terms of all the longitudinal images, and β is the penalty strength.

Observing that the DS penalties given in Section 3.1 are functions only of the difference image, here the set of alignment-corrected difference images, $\{\boldsymbol{\delta}_{sk}\}$, is defined as

$$\boldsymbol{\delta}_{sk} = \alpha_k \boldsymbol{M}_k \boldsymbol{\theta}_k - \alpha_s \boldsymbol{M}_s \boldsymbol{\theta}_s, \tag{4.2}$$

where $\boldsymbol{M}_s$ is the $J \times J$ sparse matrix form of the alignment operator for the sth longitudinal image and α_s is a scaling factor that normalises $\boldsymbol{\theta}_s$ appropriately so that voxels that are longitudinally unchanging, in terms of metabolic behaviour, produce differences of zero. Note that all images are aligned to a single common reference space, allowing the use of only S alignment operators. In the case where this reference space coincides with one of the longitudinal scans, $\boldsymbol{\theta}_s$, $\boldsymbol{M}_s$ becomes the identity matrix and only $S-1$ alignment operators must be calculated.

Given this set of difference images, the general penalty term is defined as

a summation of contributions from each difference image, i.e.

$$U(\{\boldsymbol{\theta}_s\}) = \sum_{s=1}^{S}\sum_{k=1}^{S} w_{sk}u(\boldsymbol{\delta}_{sk}). \tag{4.3}$$

Explicitly, given S scans, the S^2 difference images can be calculated and the total penalty is the weighted sum of the function $u(\cdot)$ applied to all possible difference images. The weights w_{sk} control the coupling between scans in the reconstruction; if the weights are chosen such that $w_{sk} = w_{ks} = 0$, the difference image between scans s and k has no contribution to the overall penalty (and thus has no penalty applied).

In order to use this form of penalty in the OSL-MAPEM algorithm, the gradient $\partial U/\partial\theta_{j,s}$ is required. Assuming that the weights matrix is symmetric (i.e. $w_{sk} = w_{ks}$), it can be shown (see Appendix for details) that this gradient is given by:

$$\frac{\partial U}{\partial \theta_{j,s}} = 2\alpha_s \sum_{k=1}^{S} w_{ks}\left[\boldsymbol{M}_s^T \nabla u_{ks}\right]_j \tag{4.4}$$

where $[\cdot]_j$ denotes the jth element of a vector expression, $\boldsymbol{M}_s^T$ is the transpose of the alignment matrix for scan s, and

$$\nabla u_{ks} = \left[\frac{\partial u(\boldsymbol{\delta}_{ks})}{\partial \delta_{1,ks}}, ..., \frac{\partial u(\boldsymbol{\delta}_{ks})}{\partial \delta_{J,ks}}\right]^T \tag{4.5}$$

is the gradient vector of u with respect to each of the input difference image voxel values for a given longitudinal image pair s, k. Note that by including the alignment in this manner, each of the PET datasets are reconstructed in their respective native image spaces; the only time alignment operators are applied is in the calculation of the penalty gradient terms.

There are numerous possibilities for the functional form of u. Three options are presented here, including the sparsity-based DS prior introduced in Section 3.1, that is included here for completeness.

4.2 Characteristics of Difference Images and their Corresponding Priors

In Section 3.1, difference images between longitudinal tumour response scans were assumed to be sparse, representing the case where only voxels in and around the tumour changed longitudinally. Here two more expected characteristics of difference images are introduced, namely that they have low entropy, and that they exhibit sparsity in their spatial gradients. Figure 4.1 shows an example simulated case, where increasing counts levels in a pair of longitudinal scans results in a difference image that is sparser, has lower entropy, and has a sparser spatial gradient image. Three difference-image based priors are therefore defined: the DS prior as defined in Section 3.1, updated to reflect the generalised reconstruction problem given in Equations 4.1–4.3; a prior based on the entropy of the difference image; and a prior based on the ℓ_1-norm of the spatial gradient image.

4.2.1 Difference-Image Sparsity (DS) Prior

As initially described in Section 3.1, and justified by the appearance of longitudinal change images in oncology (Figure 1.1), this prior assumes that only a small number of voxels changes from scan to scan, i.e. that the difference images are to be sparse. This is reiterated in Figure 4.1. Based on the results from Chapter 3, only the smooth ℓ_1-norm DS prior was used in this Chapter since the DS-NC prior was not observed to greatly improve image quality relative to the DS prior, and resulted in a less stable image reconstruction method.

In terms of a single difference image $\boldsymbol{\delta}$, the DS prior as given in Equation 3.3a is written as:

$$u_{\mathrm{DS}}(\boldsymbol{\delta}) = \sum_{j=1}^{J} \sqrt{\delta_j^2 + \varepsilon^2} \tag{4.6}$$

where ε is a small value that is used to ensure that the prior is differentiable at all values. The gradient of this function with respect to voxel j in the difference

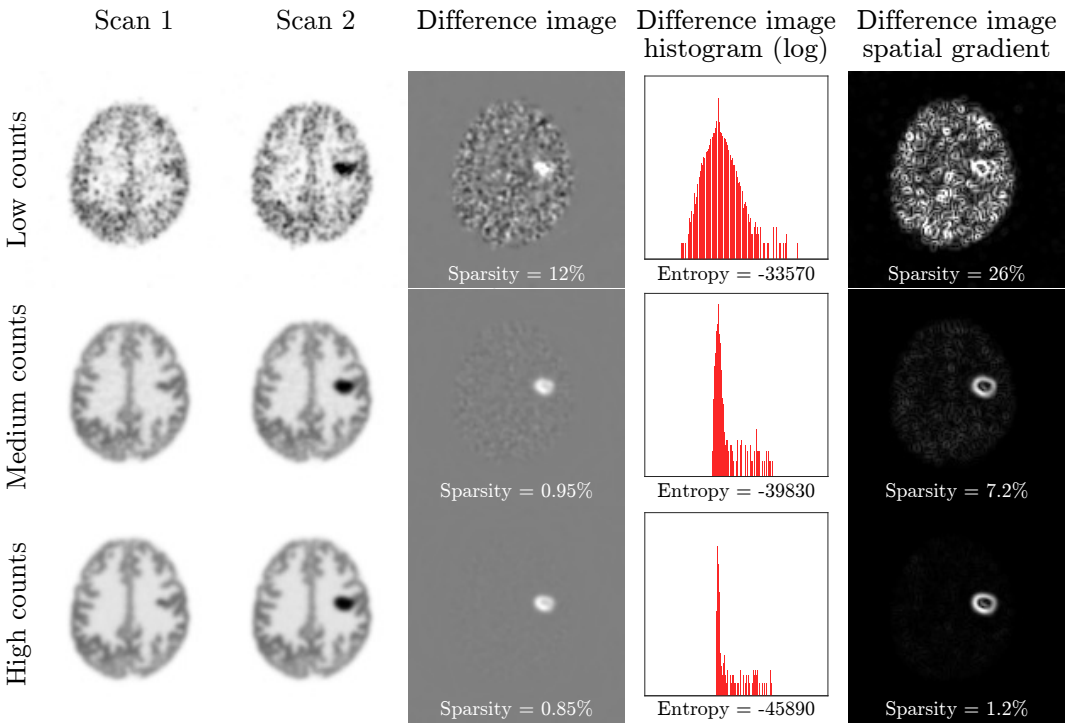


Figure 4.1: Effect of increasing counts on the characteristics of longitudinal tumour response difference images. As the counts levels increase, measures of sparsity of the difference image, entropy of the difference image, and sparsity in the spatial gradient of the difference image all decrease. This motivates the use of the DS, DE, and DTV priors respectively. Note that sparsity here was measured as the proportion of voxels with an intensity greater than 10% of the maximum value in each image and that for this reason sparsity values are only comparable within each domain and not across domains.

image $\boldsymbol{\delta}$ is given by

$$\frac{\partial u_{\mathrm{DS}}}{\partial \delta_j} = \frac{\delta_j}{\sqrt{\delta_j^2 + \varepsilon^2}}. \tag{4.7}$$

Moving from the two-scan case used in Chapter 3 to the generalised problem in Equations 4.1–4.3 simply entails the inclusion of the ℓ_1-norm of each difference image in turn, before summing the results. Note that in this generalised formulation, a longitudinal image $\boldsymbol{\theta}_s$ contributes twice to the penalty, once in $\boldsymbol{\delta}_{sk}$ and once in $\boldsymbol{\delta}_{ks}$. In the formulation used in Chapter 3, it only contributed once, since only $\boldsymbol{\phi} - \boldsymbol{\theta} = \boldsymbol{\theta}_2 - \boldsymbol{\theta}_1$ was used. This results in a change of scale for β values in this Chapter.

4.2.2 Difference-Image Entropy (DE) Prior

As shown in Figure 4.1 difference images in longitudinal oncology scans can reasonably be expected to contain a large number of zeros and a few non-zero elements of similar intensity. As image noise increases, the spread of values in the difference image also increases, merging peaks in the histogram of the difference image and increasing the entropy of the difference images. To reflect this, an entropy-based difference-image prior (difference entropy, or DE) that penalises higher entropy difference images is introduced. The use of information-based priors has previously been explored in the field of anatomically guided PET image reconstruction [168, 174–179], but to our knowledge, they have not previously been incorporated into longitudinal PET reconstructions or applied to difference images.

The PDF of a continuous random variable, X, given a set of parameters, $\{\delta_j\}$, can be estimated as $\hat{p}(x;\boldsymbol{\delta})$ using the Parzen window, or kernel density estimation, method [235, pp. 131-137]. In this work Gaussian kernels are used, so that the PDF is estimated as

$$\hat{p}(x;\boldsymbol{\delta}) = \frac{1}{J}\sum_{j=1}^{J} C\exp\left(-\frac{(x-\delta_j)^2}{2W^2}\right), \tag{4.8}$$

where W is the width of the kernel and C is a normalising scalar that sets the integral of each kernel to one. With this PDF the entropy of a difference

image $\boldsymbol{\delta}$ can be estimated according to:

$$\hat{H}(X;\boldsymbol{\delta}) = -\int \hat{p}(x;\boldsymbol{\delta}) \log \hat{p}(x;\boldsymbol{\delta}) dx. \tag{4.9}$$

As previously mentioned, the DE prior aims to encourage solutions which have a low entropy difference image. Therefore, the prior is defined in terms of a single difference image, equal to the entropy of that difference image:

$$u_{\mathrm{DE}}(\boldsymbol{\delta}) = \hat{H}(X, \boldsymbol{\delta}). \tag{4.10}$$

In practice the continuous variable X is replaced with a discrete approximation, and the integral in Equation 4.9 becomes the following sum

$$\begin{aligned} \hat{H}(X;\boldsymbol{\delta}) &\approx \hat{H}(\boldsymbol{x};\boldsymbol{\delta}) \\ &\approx -\sum_{l=1}^{N_l} \hat{p}(x_l;\boldsymbol{\delta}) \log \hat{p}(x_l;\boldsymbol{\delta}), \end{aligned} \tag{4.11}$$

where x_l is the lth sampling point of the PDF and N_l is the number of such samples. With this approximation, the gradient of the DE prior u_{DE} with respect to a given difference image voxel δ_j is given by [178]

$$\frac{\partial u_{\mathrm{DE}}}{\partial \delta_j} = -\frac{\Delta x}{J} \sum_{l=1}^{N_l} (1 + \log \hat{p}(x_l;\boldsymbol{\delta})) \left(\frac{x_l - \delta_j}{W^2} \right) \exp\left(-\frac{(x_l - \delta_j)^2}{2W^2} \right), \tag{4.12}$$

where the interval between x_l and x_{l+1} is denoted Δx. Note that this equation requires the contribution from each and every voxel δ_j to the PDF at each and every kernel centre x_l.

Similarly to the DS-NC-PML method of Chapter 3, the DE prior is non-convex, meaning that there are potentially multiple local optima of the objective function. Nonetheless, non-convex entropy-based priors have been used previously in the PET image reconstruction literature, demonstrating good results [168, 175–179].

4.2.3 Difference-Image Total Variation (DTV) Prior

As seen in Figure 4.1, images which are intrinsically sparse can also possess sparsity in the spatial gradient domain, especially when the non-zero elements in the image in question are spatially grouped. Sparsity in the gradient of the difference images can therefore be encouraged, corresponding to the total variation (TV) of the difference images. The TV of a 2D difference image $\boldsymbol{\delta}$ can be defined as a smooth ℓ_1-norm of the finite-difference gradient image [136], according to

$$\mathrm{TV}(\boldsymbol{\delta}) = \sum_{x,y} \sqrt{(\delta_{x+1,y} - \delta_{x,y})^2 + (\delta_{x,y+1} - \delta_{x,y})^2 + \varepsilon^2}, \tag{4.13}$$

where $\delta_{x,y}$ here denotes the difference image indexed in each spatial dimension separately. If the difference images are expected to have sparse spatial gradient images, as suggested in Figure 4.1, this is equivalent to the TV of the difference images being low. Therefore, the difference image TV prior (DTV) uses the following for $u(\boldsymbol{\delta})$

$$u_{\mathrm{DTV}}(\boldsymbol{\delta}) = \mathrm{TV}(\boldsymbol{\delta}), \tag{4.14}$$

with a derivative with respect to voxel $\delta_{x,y}$ given by [136]:

$$\begin{aligned}\frac{\partial u_{\mathrm{DTV}}}{\partial \delta_{x,y}} =& \frac{\delta_{x,y} - \delta_{x-1,y}}{\sqrt{(\delta_{x,y} - \delta_{x-1,y})^2 + (\delta_{x-1,y+1} - \delta_{x-1,y})^2 + \varepsilon^2}} \\ &+ \frac{\delta_{x,y} - \delta_{x,y-1}}{\sqrt{(\delta_{x+1,y-1} - \delta_{x,y-1})^2 + (\delta_{x,y} - \delta_{x,y-1})^2 + \varepsilon^2}} \\ &- \frac{\delta_{x+1,y} + \delta_{x,y+1} - 2\delta_{x,y}}{\sqrt{(\delta_{x+1,y} - \delta_{x,y})^2 + (\delta_{x,y+1} - \delta_{x,y})^2 + \varepsilon^2}}\end{aligned} \tag{4.15}$$

The TV penalty as given by Equation 4.13 has previously been used in single-dataset PET reconstruction (e.g. [136]), but to the best of the author's knowledge, it (or any other spatial TV prior) has never been applied to difference images between longitudinal PET scans.

4.3 2D Hyperparameter Characterisation

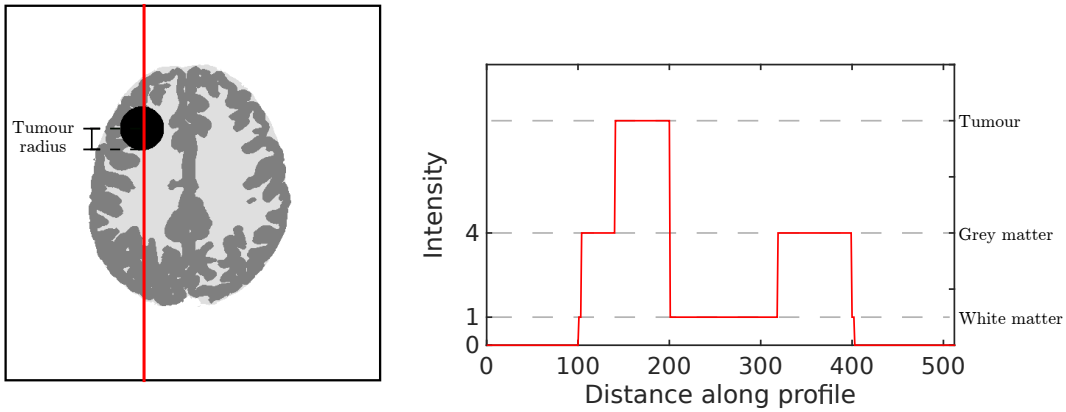
In order to characterise the effects of the DS, DE, and DTV priors, and to select β values for future 2D simulation experiments, a 2D hyperparamter characterisation study was performed. This 2D simulation experiment aimed to investigate the relative performances of the three priors without the confounding factors of misalignment and varying background activity levels.

4.3.1 Experimental Methods

Using the freely available NCAT digital brain phantom [227], realistic 2D [^{18}F]-FDG ground truth images were created, comprising four segmented intensity regions: air, white matter, grey matter, and tumour (Figure 4.2). These images were of a high resolution, with an array size of 512×512 and a pixel size of $0.5 \times 0.5\,\text{mm}^2$. Air voxels were assigned activities of 0 and the ratio between grey and white matter intensities was defined as 4:1. The tumour was a variable uniform circular region on the boundary between the left middle frontal gyrus and white matter, characterised by its radius and intensity.

The process for producing simulated PET datasets was similar to that used in Section 3.2.1. Ground truth images were blurred by a 4.3 mm FWHM Gaussian kernel before being projected into sinogram space. Attenuation factors were calculated using a corresponding attenuation map consisting of air ($\mu = 0\,\text{mm}^{-1}$), soft tissue ($\mu = 0.0096\,\text{mm}^{-1}$), and bone ($\mu = 0.0172\,\text{mm}^{-1}$) [106] voxels. Randoms were approximated as uniform sinograms, and scattered coincidences were approximated by smoothing the projections of the ground truth images with a Gaussian kernel of 24 bins FWHM. These scatters and randoms were scaled and added to the data to produce noise-free sinograms which contained 20% scatter and 20% randoms. As in Section 3.2, noise-free and non-resolution degraded datasets were reconstructed with FBP to provide reference images in the same space as the noisy reconstructions.

To explore the effect of β on reconstructed image quality, the two-scan case was investigated, with ground truth images as shown in Figure 4.3. The two



(a) (b)

Figure 4.2: The ground truth image tumour model used for 2D simulation studies. (a) An example ground truth image, and (b) the indicated profile through the image intersecting the tumour. The ground truth images comprised of four classes of pixel: air, white matter, grey matter, and tumour, with each one assigned a single intensity value. With a nominal white matter intensity of 1, the grey matter intensity was 4. The tumour was characterised by its radius and intensity. Note that ground truth images had an array size of 512×512 and a voxel size of $0.5 \times 0.5\,\text{mm}^2$.

ground truth images were generated as outlined above, one baseline (PET1) and one follow-up (PET2). The PET1 tumour had a radius of 15 mm and a tumour to white matter intensity ratio of 8:1 (Figure 4.3(a)), and the PET2 tumour was 8 mm with a corresponding ratio of 6:1 (Figure 4.3(b)). The two images were scaled with a common factor such that the total expected number of counts in the PET1 sinogram was ∼4 million. Finally, a double-counts pair of noise-free datasets were generated to provide low-noise benchmark images, as in Section 3.2.

100 noisy realisations of each of the datasets (PET1, PET2, double-counts PET1, and double-counts PET2) were then produced by introducing Poisson noise into the noise-free datasets. These datasets were reconstructed with a variety of methods, all run for 200 iterations, and all including the original attenuation, scatters and randoms sinograms from the simulation process in the forward model. Resolution modelling was not included to remain consistent with the results presented in Chapter 3. The double-counts PET1 and PET2 datasets were reconstructed with MLEM only in order to demonstrate the image quality corresponding to doubling the number of acquired counts. The

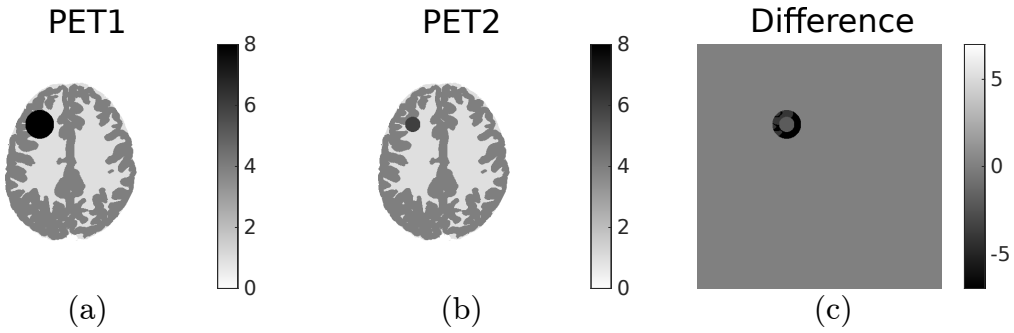


Figure 4.3: Ground truth images for the 2D hyperparameter experiment. The PET1 (a) tumour had a radius of 15 mm and an intensity of 8 (arb. units), and the PET2 (b) tumour was 8 mm in radius with an intensity of 6. (c) The resulting ground truth difference image.

remaining datasets were reconstructed with MLEM, DS-PML, DE-PML, and DTV-PML as follows.

All elements of the scan-to-scan weighting matrix were set to 1, as were all activity correction factors α_s. The alignment matrices $\boldsymbol{M}_s$ were set to the identity matrix $\boldsymbol{I}$, since ground truth images were generated with perfect alignment. The prior-specific parameters were as follows:

DS-PML: β values were varied between 7.5×10^{-5} and 15×10^{-4} and ε was set to 1×10^{-6}.

DE-PML: A Gaussian window of standard deviation $W = 0.35$ (arb. intensity units) was used to estimate the PDF of the difference images. During each iteration of the DE-PML algorithm, a 100-level discretisation was used for approximating the continuous PDF. In addition, a mask was used to ensure that entropy was only calculated in the brain itself since the remaining air voxels naturally exhibit low entropy. β values were varied between 0.15 and 3.

DTV-PML: β values were varied between 4×10^{-5} to 8×10^{-4} and ε was set to 1×10^{-6}.

In addition, the longitudinal smooth discussed in Section 3.4 was applied to the MLEM reconstructions, using a two-point Gaussian kernel parameterised by the standard deviation, σ, so that with $\sigma = 0$ the images were unchanged, and with $\sigma = +\infty$ the output images were the unweighted average of the input images.

To evaluate the performance of the various reconstruction methods, tumour

%RMSE values (as defined in Equation 3.5) were calculated for PET1 and PET2 images using the non-resolution-degraded, noise-free FBP images as the corresponding reference images, as in Section 3.2. The analysis region, $\mathcal{T}$, was selected as a 226-voxel circular ROI covering the PET1 15 mm tumour and a small portion of surrounding tissue. Note that this tumour region was used in both PET1 and PET2 images so that the PET2 image contained more background within $\mathcal{T}$. In addition, aggregated CV was calculated according to:

$$\mathrm{CV}_\Omega(\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_N) = \frac{1}{N}\sum_{n=1}^{N}\frac{\mathrm{SD}_\Omega(\boldsymbol{\theta}_n)}{\mathrm{Mean}_\Omega(\boldsymbol{\theta}_n)} \tag{4.16}$$

where n indexes one of N noisy reconstructed images, and Ω is the analysis region (c.f. Equation 3.6). CV was calculated in an eroded mask of all white matter voxels to provide a measure of image noise. Since Chapter 3 showed that the reduction in whole-brain %RMSE was primarily due to noise reduction, white matter CV serves as a metric of background image quality in this Chapter.

The mean and CV of voxel intensities within the PET1 and PET2 tumours (*excluding* surrounding tissue voxels) were also calculated for each reconstruction. The average of these values across all noisy realisations was calculated to provide a measure of the effects of the regularisation penalties on tumour quantification. This is equivalent to measuring the effect of the proposed penalties on the SUV within the tumour, with an arbitrary scaling factor.

Finally, bias and standard deviation (SD) maps were produced by calculating the voxel-wise bias and standard deviation of the 100 noisy realisations of the data i.e.:

$$\mathrm{Bias}_j = \bar{\theta}_j - \theta_j^{\mathrm{Ref}} \tag{4.17a}$$

$$\mathrm{SD}_j = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}\left(\theta_{jn} - \bar{\theta}_j\right)^2} \tag{4.17b}$$

where $\bar{\theta}_j$ is the mean reconstructed image over all noisy realisations. The bias and SD maps allow for visual inspection of the spatial distribution of reconstruction error, allowing qualitative comparison of the proposed methods.

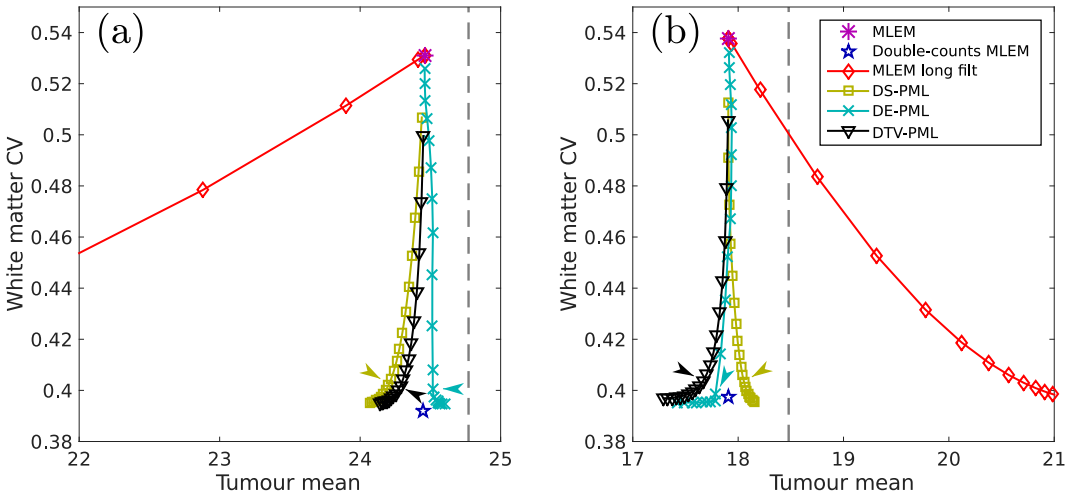


Figure 4.4: Tumour mean vs white matter CV as a function of regularisation hyperparameter for (a) PET1 and (b) PET2. For the longitudinally smoothed MLEM reconstructions (red diamonds), the hyperparameter is σ, controlling the width of the filter, and for DS-PML (yellow squares), DE-PML (cyan crosses) and DTV-PML (black triangles), the parameter is β. See text for the ranges of σ and β values used for each method. Also shown are the performances of MLEM (purple asterisks) and double-counts MLEM (blue stars). Vertical dashed lines show the tumour mean values in the non-resolution-degraded, noise-free FBP reference images. Arrows indicate the selected β values for each method, as listed in Table 4.1.

4.3.2 Results

Figure 4.4 shows the PET1 and PET2 tradeoffs between tumour mean and white matter CV as a function of regularisation hyperparameter for the proposed methods. In this context, an ideal method would provide a point that coincides with the double-counts MLEM point using the same counts level available to the proposed methods.

The longitudinal smooth, DS-PML, DE-PML, and DTV-PML reconstructions all reduce white matter CV down to levels similar to double-counts MLEM when sufficiently high hyperparameter values are used, although at the cost of varying levels of bias in the tumour mean. In terms of limiting this bias, it is clear that the proposed methods are far superior to the weighted average longitudinal smooth. In particular, the DE-PML method was observed to produce the lowest tumour bias while allowing noise reduction approaching double-counts MLEM.

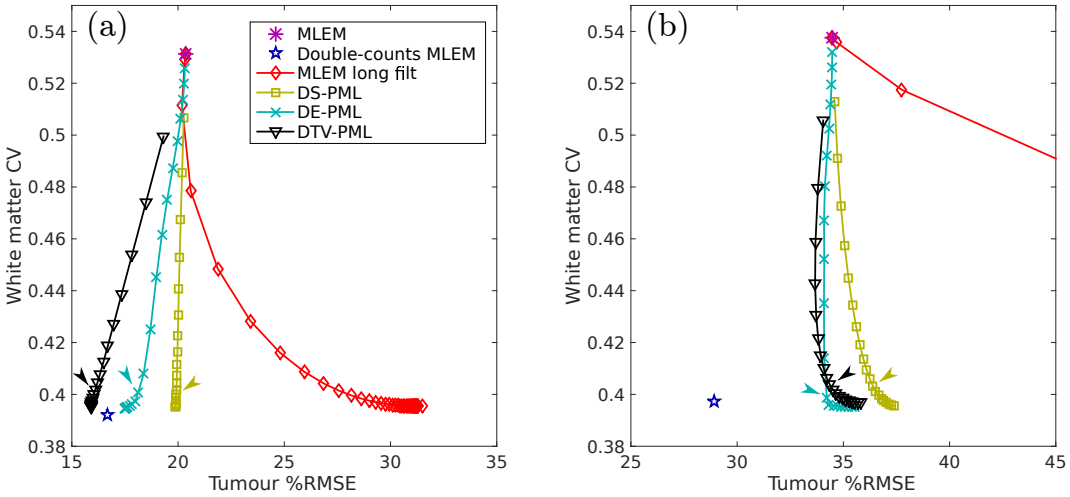


Figure 4.5: Tumour %RMSE vs white matter CV as a function of regularisation hyperparameter for (a) PET1 and (b) PET2 reconstructions. Reconstruction methods shown are MLEM, DS-PML, DE-PML, and DTV-PML. Also shown are longitudinally smoothed MLEM and double-counts MLEM reconstructions. Arrows indicate the selected β values listed in Table 4.1.

The corresponding tradeoff between tumour %RMSE and white matter CV is shown in Figure 4.5. Compared to the MLEM reconstructions, the double-counts MLEM reconstructions exhibit lower tumour %RMSE in addition to lower white matter CV. While the DS-PML method slightly reduced PET1 tumour error compared to MLEM (similarly to as seen previously in Figure 3.4 in Chapter 3), the DE-PML and the DTV-PML methods reduce PET1 tumour error to a greater extent. Furthermore, the DTV-PML method with high β values reduces error slightly more than the double-counts MLEM reconstruction. For the PET2 reconstructions, the three difference-image priors produce tumour %RMSE values more similar to the MLEM reconstructions, increasing at high β values as the penalty strength becomes too large. The DE-PML and DTV-PML methods produce lower tumour %RMSE values than DS-PML.

To assess the tradeoff between bias and variance that gives rise to the %RMSE values in Figure 4.5, the tumour mean was plotted against the tumour CV as a function of β for the presented difference-image prior methods, as shown in Figure 4.6. It can be observed that both DTV-PML and DE-PML produced noise reduction in the tumour itself, resulting in the lower %RMSE

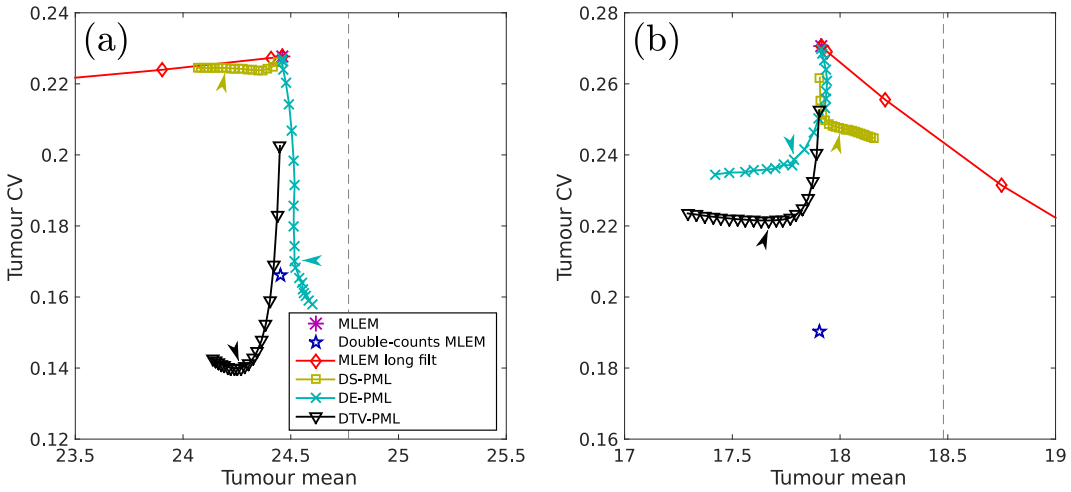


Figure 4.6: Tumour mean against tumour CV for varying regularisation parameter for (a) PET1 and (b) PET2 reconstructions. DS-PML, DE-PML and DTV-PML methods are shown as a function of β. Also shown are MLEM, MLEM with double-counts data, and longitudinally filtered MLEM, parameterised by the width of the filter, σ. Arrows indicate the β values listed in Table 4.1, used in the remainder of the 2D simulation experiments.

values observed previously. The DS-PML method performed worse than the other two methods in terms of noise reduction. It is interesting to note that in the PET1 reconstructions, DE-PML and DTV-PML reduced tumour noise to levels below even the double-counts reconstruction when high values of β were used. For the PET2 reconstructions, tumour noise plateaued at a CV level greater than the double-counts MLEM case.

Based on Figures 4.4–4.6, the selection of an optimal β value for each method is non-trivial. In practice, there is a tradeoff between all metrics investigated, so that any single β value cannot be optimal for all measures of image quality. Since optimising so many correlated and anti-correlated metrics simultaneously is difficult, and since ultimately any clinical usage of these algorithms would be subject to the requirements of the particular application in question, selection of β values based on a quantitative measure was not attempted. Instead, to demonstrate both the full potential offered by each method and their respective drawbacks, β values were selected so as to approach the noise levels of the double-counts MLEM reconstructions, without excessive penalisation. The selected values are indicated by arrows in Figures 4.4–4.6, and listed in Table 4.1.

In order to verify that the trends observed in Figures 4.4–4.6 are due to the form of DS, DE, and DTV priors, rather than their convergence when included in the OSL algorithm, a single noisy realisation of the datasets was reconstructed with each prior using the corresponding β values listed in Table 4.1. The data was also reconstructed with MLEM. These reconstructions were run to 1000 iterations, and the value of the objective function for each method at each iteration number was calculated. Figure 4.7 shows the normalised objective function as a function of iteration number for MLEM and the three difference-image prior methods. Observing the progression of the normalised objective function with iteration number, it is apparent that all of the methods converge with a similar trajectory to MLEM, with each one within 99.99% of their 1000-iteration value at 200 iterations. This suggests that the comparisons made in Figures 4.4–4.6 were fair, with performances representative of the methods at convergence.

To put into perspective the bias levels shown in Figure 4.4, Figure 4.8 plots the tumour mean values for the proposed methods with the selected β values as listed in Table 4.1. The MLEM and double-counts MLEM tumour means are also shown for comparison. The biases previously observed in Figure 4.4 are apparent, but the effect is small relative to the magnitude of the values involved, even for the most severe cases (PET2 DTV-PML or both DS-PML reconstructions).

When considering the bias and standard deviation maps (Figure 4.9), it is possible to assess the spatial distribution of error for the proposed methods using the β values given in Table 4.1. It is apparent that all of the proposed methods reduce the voxel-wise variance in background regions due to the noise reduction indicated by Figures 4.4 and 4.5. However, behaviour in and around

Method	Selected β value
DS-PML	9×10^{-4}
DE-PML	1.8
DTV-PML	4×10^{-4}

Table 4.1: β values selected for as representative of the performances of DS-PML, DE-PML, and DTV-PML based on Figures 4.4–4.6

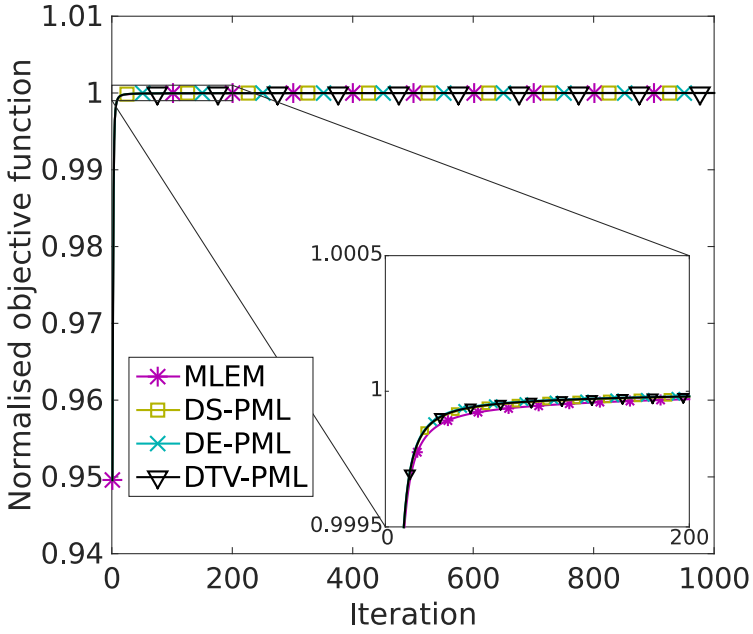


Figure 4.7: Plot of the objective function as a function of iteration number for a single noisy realisation of the experiment 1 datasets reconstructed with MLEM, DS-PML, DE-PML, and DTV-PML. For each of the PML reconstructions, β was set to the value indicated in Table 4.1. Note that for each method, the objective function values have been normalised by the 1000-iteration value.

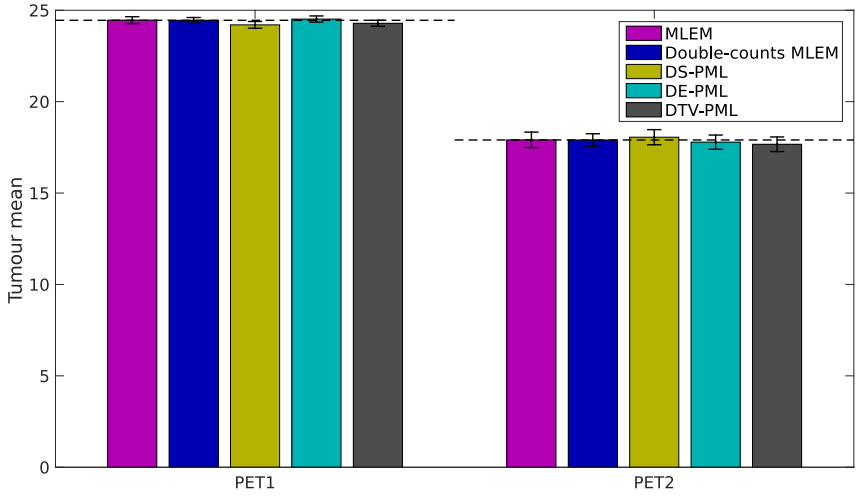


Figure 4.8: Tumour mean values for MLEM, double-counts MLEM, DS-PML, DE-PML and DTV-PML, all at 200 iterations. β values for the PML methods are those listed in Table 4.1. Error bars show the standard deviation of the measured mean values over the 100 noisy realisations of the data. Dashed lines replicate the double-counts MLEM tumour means for comparison.

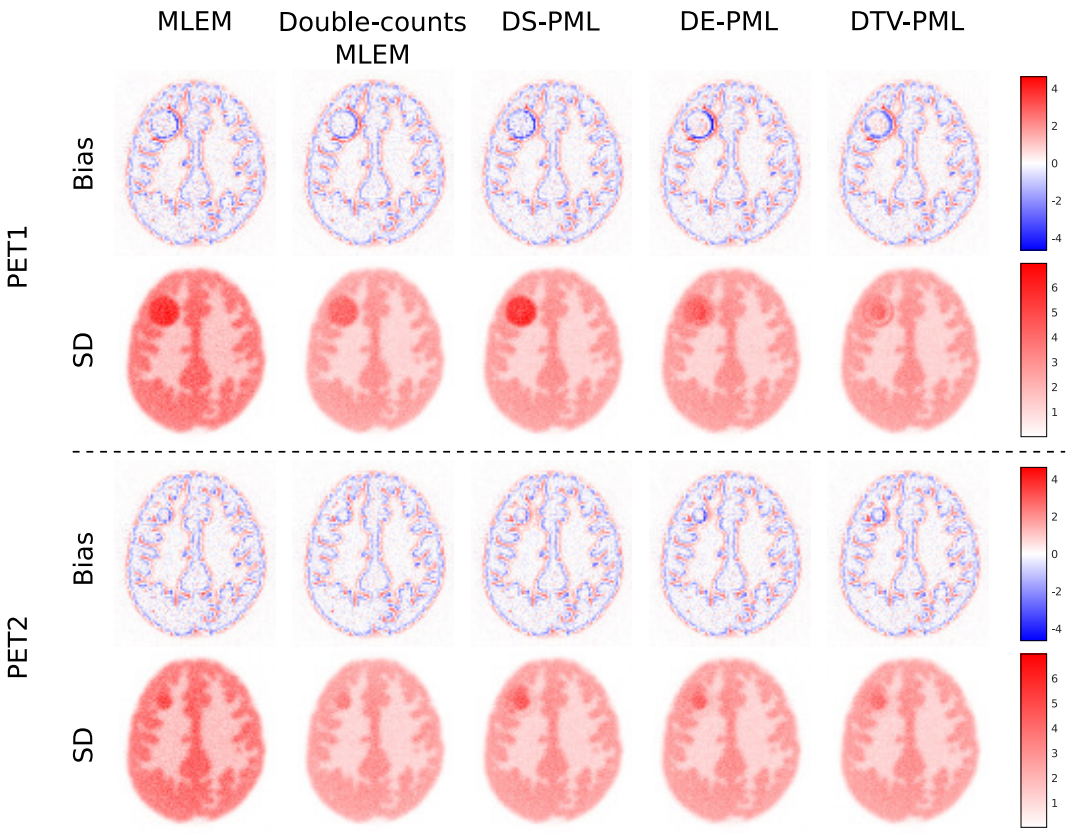


Figure 4.9: Bias and standard deviation (SD) images calculated over 100 noisy realisations for the PET1 and PET2 datasets reconstructed with MLEM, double-counts MLEM, DS-PML, DE-PML and DTV-PML. β values for the PML methods are the same as those listed in Table 4.1.

the tumour is different for the various methods, with DS-PML maintaining voxel-wise variance at levels similar to MLEM, and with the DE-PML and DTV-PML methods reducing voxel-wise variance in the tumour (c.f. Figure 4.6). In terms of the bias maps, the difference-image prior methods produce similar results to the MLEM reconstructions in the background of the image, but with differing behaviour in the tumour. The DS-PML method exhibits a visible negative bias (blue) within the PET1 tumour compared to MLEM, with DE-PML and DTV-PML introducing higher levels of bias around the edge of the tumour. In the PET2 tumour, the bias around the edges of the tumour using DE-PML and DTV-PML is more apparent.

4.4 Effect of Tumour Change on Reconstruction Quality

In order to evaluate the effect of changing the penalty strengths using the proposed methods, the previous Section considered only one example of a tumour response. In practice, the performance of these methods, like any reconstruction method, will be dependent on the object being imaged. Therefore, different tumours may affect reconstruction performance. In cancer imaging, where the size and intensity of tumours cannot be known *a priori*, it is important to validate reconstruction methods over a range of tumour characteristics. This is particularly true for the proposed methods, since they rely on assumptions about the nature of the difference image, the validity of which depends on the tumour changes encountered.

4.4.1 Experimental Methods

To evaluate the performance of the difference-image prior methods for a range of longitudinal tumour behaviours, tumour radius and intensity, as described in Figure 4.2, were varied and the performance of the reconstruction methods with β values from Table 4.1 was assessed.

Initially keeping the tumour intensities the same as listed in Section 4.3.1 (8 and 6 for PET1 and PET2 respectively), the PET1 and PET2 tumour radii were varied between 3 mm and 15 mm in steps of 2 mm. 100 noisy realisations of simulated PET data were generated for each radius pair using the same methodology and counts levels as in Section 4.3, and reconstructed with 200 iterations of each method. For all PET1 and PET2 reconstructions for all tumour pairs, %RMSE was calculated in the same 226-voxel circular ROI used in Section 4.3.

Following this, the radii of the two tumours were kept the same as listed in Section 4.3.1 (15 mm and 8 mm), with the intensities of each varied between 2 and 8. Using 200 iterations of each reconstruction method, with 100 noisy

realisations per dataset again, %RMSE was calculated in the same tumour ROI.

4.4.2 Results

Figure 4.10 displays the %RMSE in the tumour region as a function of PET1 and PET2 tumour radius, and Figure 4.11 shows the corresponding results for changing PET1 and PET2 tumour intensities. The CV values in background regions were similar to those shown in Figures 4.4 and 4.5 and so are not replicated here.

Firstly, it is apparent that for all tumour sizes doubling the number of counts used in the MLEM reconstruction reduces reconstruction errors, reflecting the results shown in Figure 4.5. In general, a smaller tumour radius and intensity causes higher reconstruction errors in MLEM reconstructions.

With varying tumour radii (Figure 4.10), error levels are generally the same as or reduced compared to MLEM reconstructions when using the difference-image priors. With the difference-image prior methods, the highest errors occur when the change in tumour radius is greatest (c.f. bottom right and top left corners of the heat maps shown in Figure 4.10). For the DS-PML method, these extreme cases have %RMSE levels similar to the MLEM reconstructions, while elsewhere %RMSE is reduced slightly due to noise reduction. For the DE-PML and DTV-PML methods, even at the most extreme radii changes, error levels are lower than the MLEM case, in agreement with Figure 4.5. In some cases these errors approach the double-counts MLEM reconstructions.

For the reconstruction errors in the tumour region as a function of varying tumour intensties (Figure 4.11), similar results are observed, with the DS-PML error levels similar to the MLEM errors, and the DE-PML and DTV-PML errors falling between the MLEM and the double-counts MLEM errors.

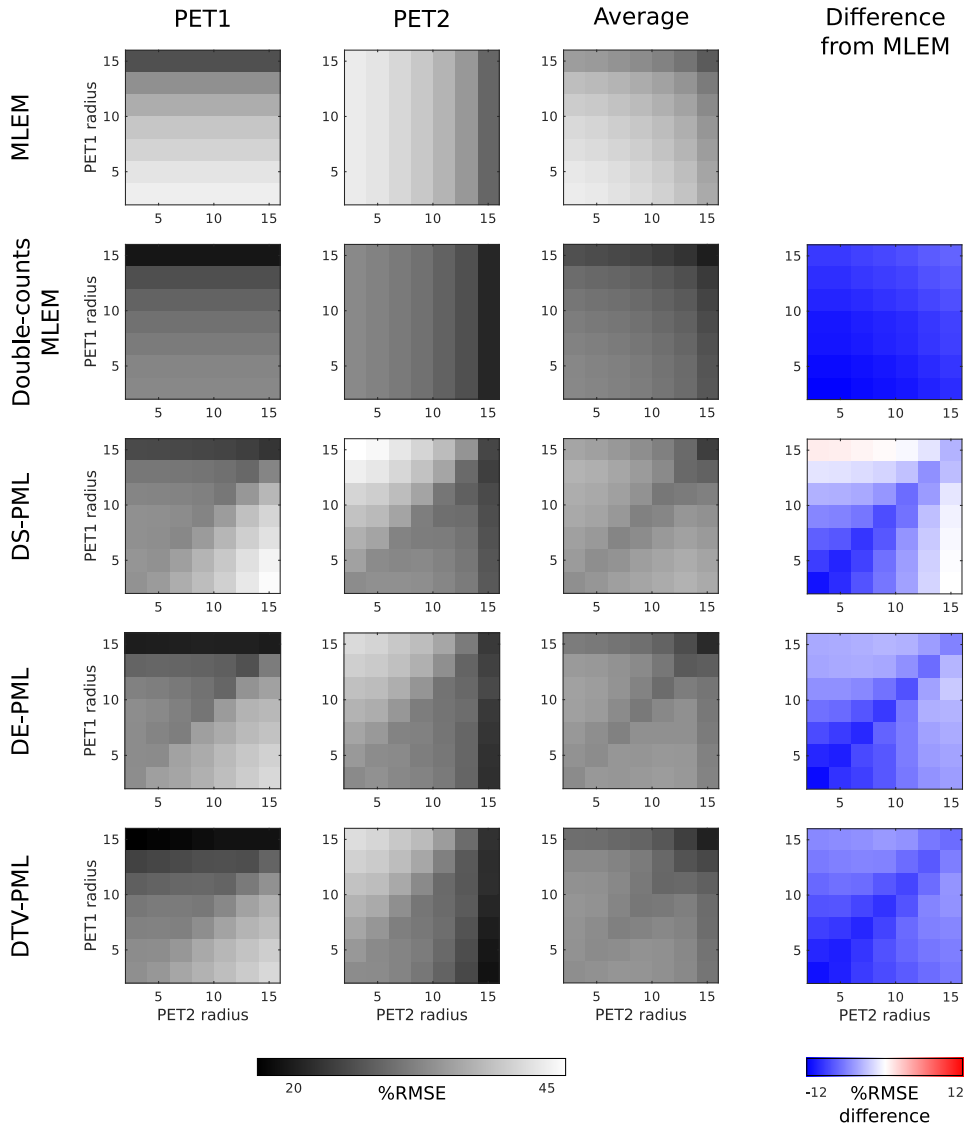


Figure 4.10: %RMSE values in PET1 and PET2 reconstructions in a region encompassing the tumour for varying tumour radii for: MLEM, double-counts MLEM, and DS-PML, DE-PML and DTV-PML with β values from Table 4.1. Tumour intensities were kept constant at 8 and 6 (relative to a nominal white matter uptake of 1). Also shown are average %RMSEs across PET1 and PET2 and the difference of these averages from the MLEM averages. Note that in all cases the analysis region was a dilation of the largest (15 mm) tumour, and so for smaller tumours the analysis region contained proportionally more background. It is apparent that across tumour sizes, the difference-image prior methods allow for good reconstruction of the region around the tumour, with error values generally falling between MLEM and double-counts MLEM reconstructions.

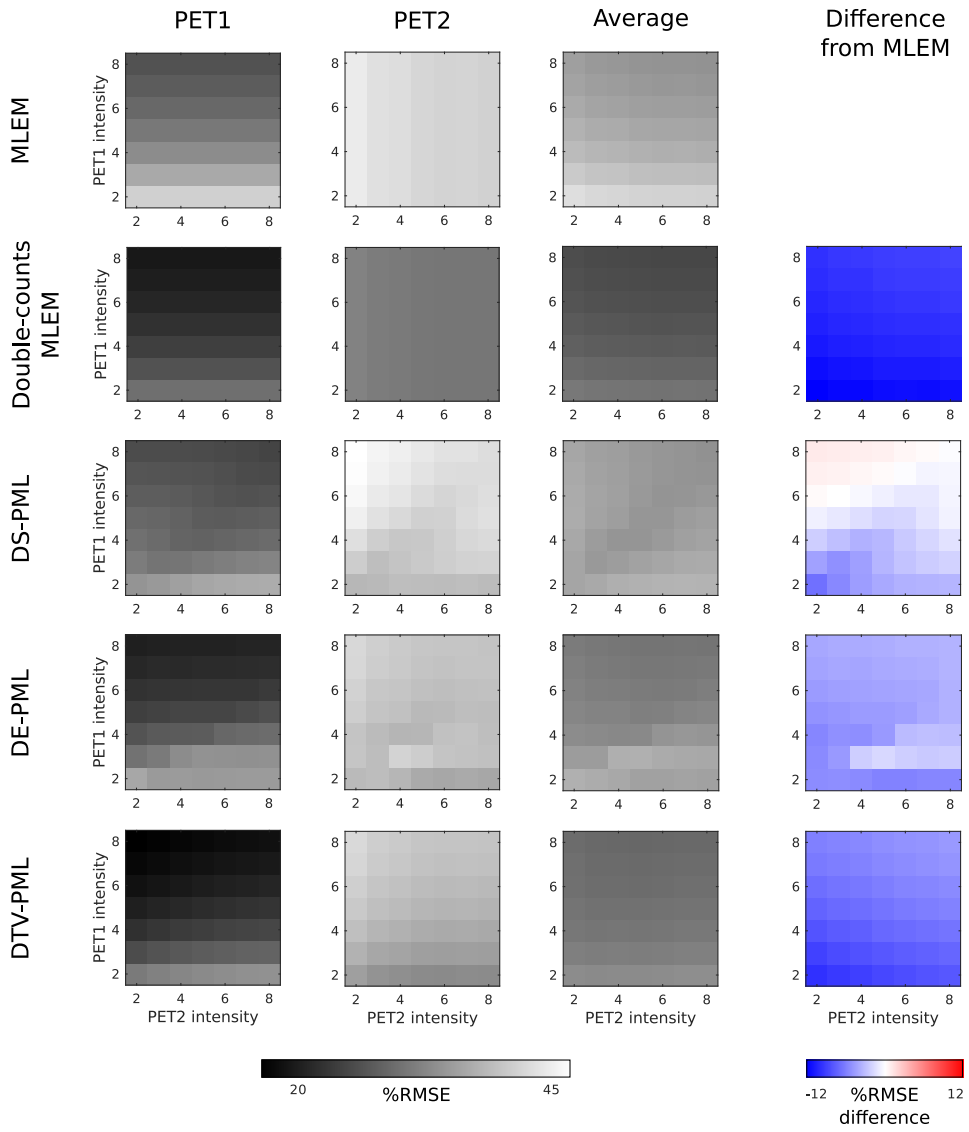


Figure 4.11: %RMSE values in PET1 and PET2 reconstructions in a region encompassing the tumour for varying tumour intensities for: MLEM, double-counts MLEM, and DS-PML, DE-PML and DTV-PML with β values as listed in Table 4.1. Tumour radii were kept constant at 15 mm and 8 mm. Also shown are average %RMSEs across PET1 and PET2 and the difference of these averages from the MLEM values. Results are similar to those seen in Figure 4.10, with error levels generally between MLEM and double-counts MLEM, although both the DS-PML and DE-PML methods produce %RMSE values more similar to MLEM for a greater number of intensity changes.

4.5 2D Five-Scan Experiment

The previous experiments aimed to explore the effects of the hyperparameters and tumour behaviour on the reconstructed image quality using the proposed methods. To achieve this, experiments using only two scans were performed to simplify the interpretation of the results. However, it is expected that the proposed methods should exhibit greater levels of improvement when more datasets are available, since there is a greater total amount of data to be shared between scans. Furthermore, since performing multiple scans in an oncology treatment monitoring protocol is not uncommon (e.g. [66, 67, 73]), it is useful to understand how the proposed methods perform in these cases. The generalised reconstruction prior in Equation 4.3 contains the weighting array $\{w_{sk}\}$, which allows the coupling between scans to be determined. Here the performance of the proposed methods in a 2D simulated five-scan case is investigated, and the effects of varying these inter-scan weights on image quality is explored.

4.5.1 Experimental Methods

Five ground truth images and datasets were created in accordance with Section 4.3.1. The tumour radii were: 15 mm, 8 mm, 4 mm, 3 mm, and 5 mm, and the respective nominal intensities were 8, 6, 4, 3, and 6. The counts level of each scan was the same as that previously used, i.e. approximately 4 million counts per scan.

The five datasets were reconstructed with each of the difference-image prior methods, as described in Section 4.3.1. In addition, the w_{sk} array is now redefined so that:

$$w_{sk} = \kappa_s \exp\left(-\frac{(\Delta_{sk})^2}{2\sigma_w^2}\right) \tag{4.18}$$

where σ_w is a shape parameter than can be varied from 0 to $+\infty$. The former case gives $w_{sk} = 1$ for $s = k$ only, indicating a complete separation of the longitudinal datasets, i.e. reverting to a maximum likelihood esti-

mate. The latter case gives a uniform weighting array (i.e. $w_{sk} = w \ \forall s, k$), applying maximum penalty between all pairs of scans in the series. Δ_{sk} is the cyclic longitudinal distance between scans s and k, given by $\Delta_{sk} = \min(|s-k|, |s-k-S|, |s-k+S|)$. Using a cyclic distance results in all scans having the same total weighting with other scans, ensuring that, for example, scans 1 and 5 are not regularised less just by virtue of having no preceding or succeeding scan. κ_s is a normalising factor used to control the sum of the rows of the weights array. In this experiment κ_s was chosen so that $\sum_k w_{sk} = 2$, in agreement with the total weight used in the previous experiments where $\{w_{sk}\}$ was $\mathbf{1}_{2\times 2}$.

It is important to note that the degree to which coupling the first and last scans in the manner described above is desirable depends on the context in question. For example, in applications where the entire longitudinal examination takes place in a relatively short space of time, the first and last scans may not be expected to vary considerably, allowing cyclical coupling in this manner. On the other hand, where the first and last scans are expected to be considerably different, coupling them may be undesirable. In this experiment the assumption is made that it is acceptable to couple scans cyclically, since the changes were known *a priori*. Future work would be required to investigate the optimal coupling between longitudinal scans in specific real applications.

To investigate the effect of inter-scan coupling on reconstructed image quality, the σ_w parameter was varied between 0.4 and 8, with a final σ_w of ∞ also included to provide a case comparable to those used in Sections 4.3 and 4.4. β values for the difference-image priors were set to the values given previously in Table 4.1.

In order to provide benchmark images at different image noise levels, the five noise-free longitudinal datasets were scaled to different counts levels and reconstructed with 200 iterations of MLEM. These counts levels ranged from 1 to 5 times the original level, to correspond to the maximum expected background noise reduction from the DS-PML, DE-PML and DTV-PML methods, as suggested by Section 4.3.

100 noisy realisations of the five-scan data series were generated and recon-

structed. Reconstructed images for all methods were analysed by calculating tumour mean, tumour CV, and white matter CV (CV values were calculated using the aggregate CV of Equation 4.16). In addition, mean squared error (MSE) maps were calculated as

$$\mathrm{MSE}_j = \mathrm{Bias}_j^2 + \mathrm{SD}_j^2 \tag{4.19}$$

where the bias and SD maps are as defined in Equation 4.17.

4.5.2 Results

Figure 4.12 shows the tumour mean vs white matter CV as a function of σ_w for DS-PML, DE-PML and DTV-PML with β values as listed in Table 4.1. Also shown are the corresponding MLEM curves, with the counts level between 1 and 5 times the number used for the proposed methods. For the proposed methods, as σ_w tends towards infinity, the noise is reduced until at $\sigma_w = \infty$, the noise levels reach a minimum which is determined by the selected β value. Out of the three priors, the DE-PML method gives the least biased tumour mean values, with $\sigma_w = \infty$ providing values comparable to the MLEM reconstructions.

Observation of the PET1 and PET2 tumour mean/CV tradeoffs (Figure 4.13) reflects the results seen in Figure 4.6, despite the change of regularisation parameter from β to σ_w. In Figure 4.13, increasing σ_w reduces tumour CV for the DTV-PML method, and to a lesser extent the DE-PML method. DS-PML has a weaker effect on tumour CV; in the PET1 reconstructions, the tumour noise was maintained at levels similar to the $1\times$ counts MLEM reconstructions, despite the introduction of a bias lowering the measured tumour mean value. Of the three difference-image prior methods, DE-PML provides on average the best tradeoff between noise reduction in the background, noise reduction in the tumour and bias-mitigation in the tumour. It is important to note that the tumour biases are limited by the β values used; setting $\sigma_w = \infty$ with higher β values will induce larger biases.

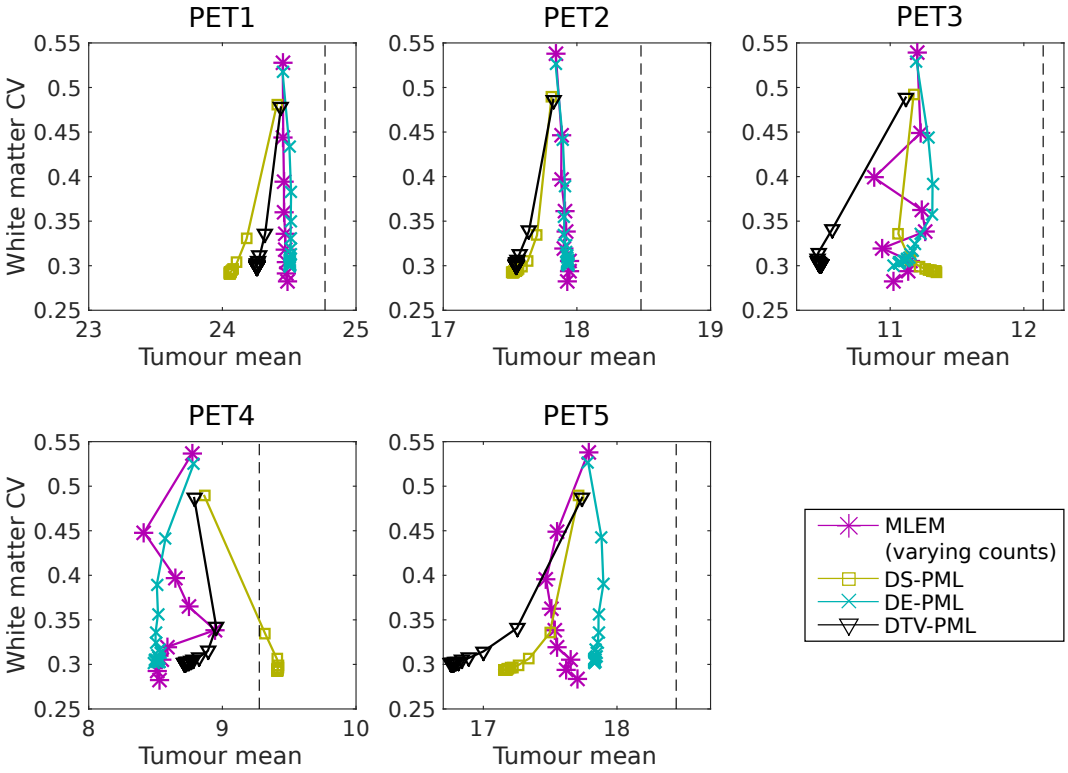


Figure 4.12: Tumour mean vs white matter CV for the five scans for DS-PML, DE-PML and DTV-PML as a function of σ_w. Penalty strengths, β, were set to the values shown in Table 4.1. Also shown is MLEM as a function of recorded counts, with the number of counts varied from 1 to 5 times the number of counts used in the proposed methods (in intervals of 0.5×). Dashed lines show the mean tumour values in the non-resolution-degraded, noise-free FBP reference images, corresponding to the nominal ground truth values of 8, 6, 4, 3, and 6 times the white matter intensity for each scan respectively. The proposed methods are able to reduce the background noise considerably, at the cost of some level of bias in the reconstructions. This bias is smallest for the DE-PML method, which remains close to the high-counts MLEM reconstructions.

When considering the example reconstructed images at $\sigma_w = \infty$ (Figure 4.14), the improvements in image quality using the proposed methods are clearly visible. Compared to the MLEM reconstructions with the same level of counts, the difference-image priors produce visually superior images with reduced noise throughout the image background, as indicated by Figure 4.12. These background noise levels are visually similar to the MLEM reconstructions with 5 times the data. However, the different effects of the proposed methods are noticeable in the tumours, particularly the PET1 tumour. The DS-PML reconstruction produces tumours that appear similar to the MLEM reconstruction with the same counts level, due to the tendency of this method

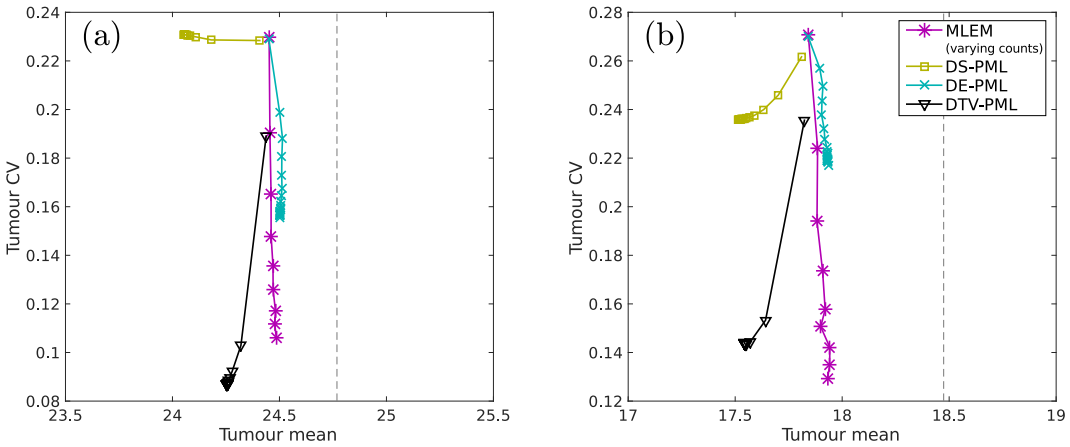


Figure 4.13: Tumour mean vs tumour CV for (a) PET1 and (b) PET2 reconstructions for the five-scan simulation experiment (PET3 to PET5 tumours have been omitted due to their small size prohibiting useful measures of tumour CV). For DS-PML, DE-PML and DTV-PML, β values are as listed in Table 4.1, and σ_w was varied between 0 and ∞. For MLEM, the number of counts per dataset was varied from 1 to 5 times the number used in the other reconstructions. For the DE-PML and DTV-PML methods, increasing σ_w reduced tumour noise while introducing some bias relative to MLEM reconstructions, whereas with DS-PML, particularly in PET1, the effect on tumour CV was weak. Dashed lines show the mean tumour values in the reference non-resolution-degraded, noise-free FBP images, corresponding to the nominal ground truth values of 8 and 6 times the white matter intensity for PET1 and PET2 respectively.

to preserve voxel-wise changes. For the DE-PML and DTV-PML methods, the tumour itself is visually improved compared to MLEM with the same number of counts, due to the noise reduction in the tumour observed in Figure 4.13.

Finally, Figure 4.15 shows the MSE images for each of the methods. Compared to MLEM reconstructions with the same amount of data, the difference-image prior reconstructions reduce MSE throughout the image background due to noise-reduction in these regions. In the tumour the levels of error reduction vary between the proposed methods. The DS-PML method slightly reduces error compared to the MLEM reconstructions, due to the small amount of noise-reduction that occurs when voxel-values are longitudinally biased offsetting the bias itself, as previously observed in Chapter 3. For the DE-PML and DTV-PML methods, error reduction in the tumour is stronger. Here it can be seen that despite the correct estimation of tumour means provided by the DE-PML method in Figure 4.12, this method still produces higher voxel-wise

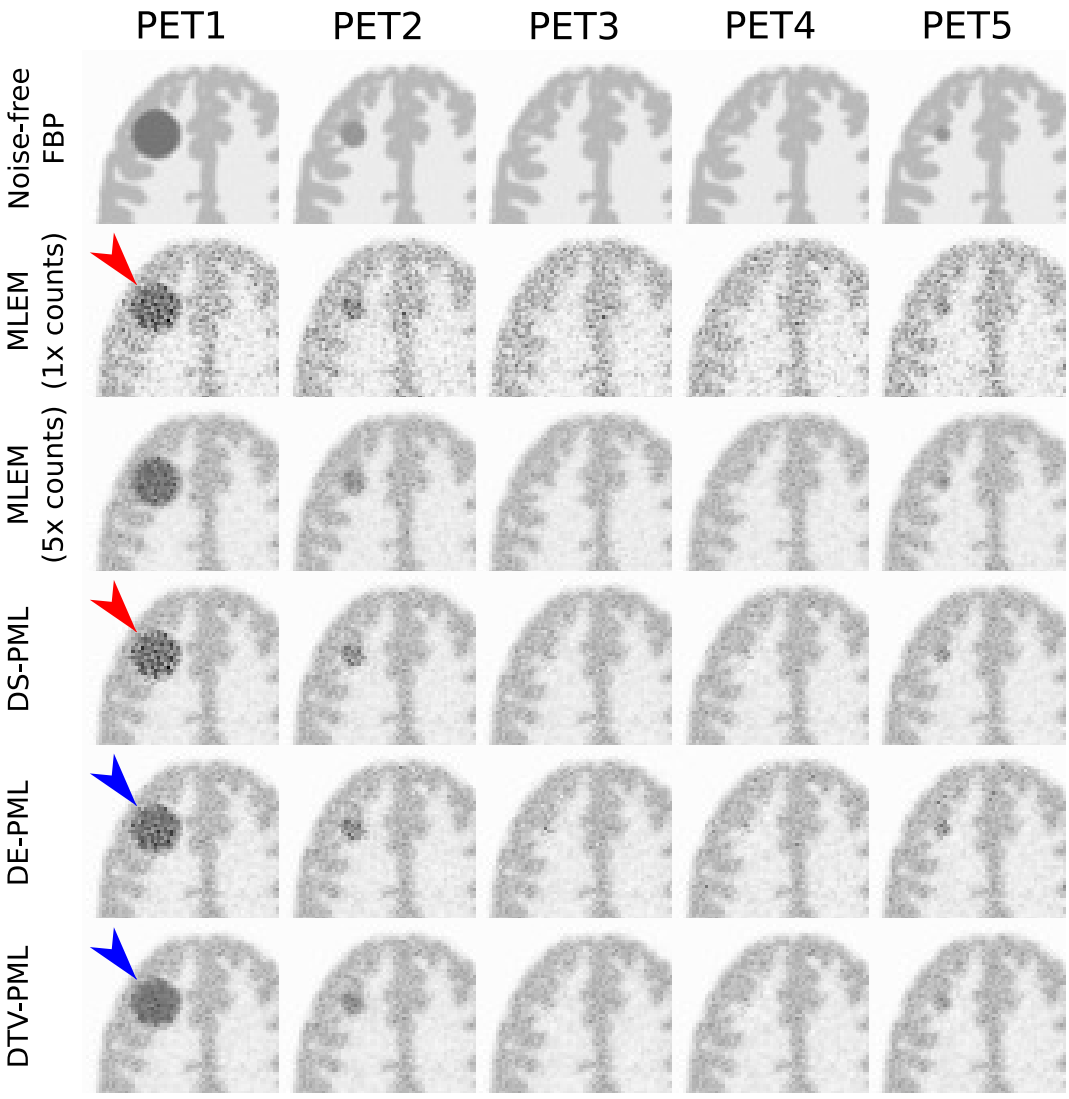


Figure 4.14: Example reconstructed images for the simulated five-scan reconstruction experiment. Each column shows the images from each of the scans. From top to bottom: the noise-free FBP reference image, MLEM with a counts level of 1, MLEM with 5 times the number of counts, DS-PML, DE-PML and DTV-PML. All of the difference-image prior reconstruction methods had $\sigma_w = \infty$ and β values as listed in Table 4.1. The three difference-image priors reduce noise in the background of the image to levels equivalent to the 5× counts MLEM, but differ in the appearance of the tumours, particularly in PET1, where DS-PML retains a similar tumour noise pattern to the 1× counts MLEM reconstruction (red arrows), whereas DE-PML and DTV-PML reduce the noise (blue arrows).

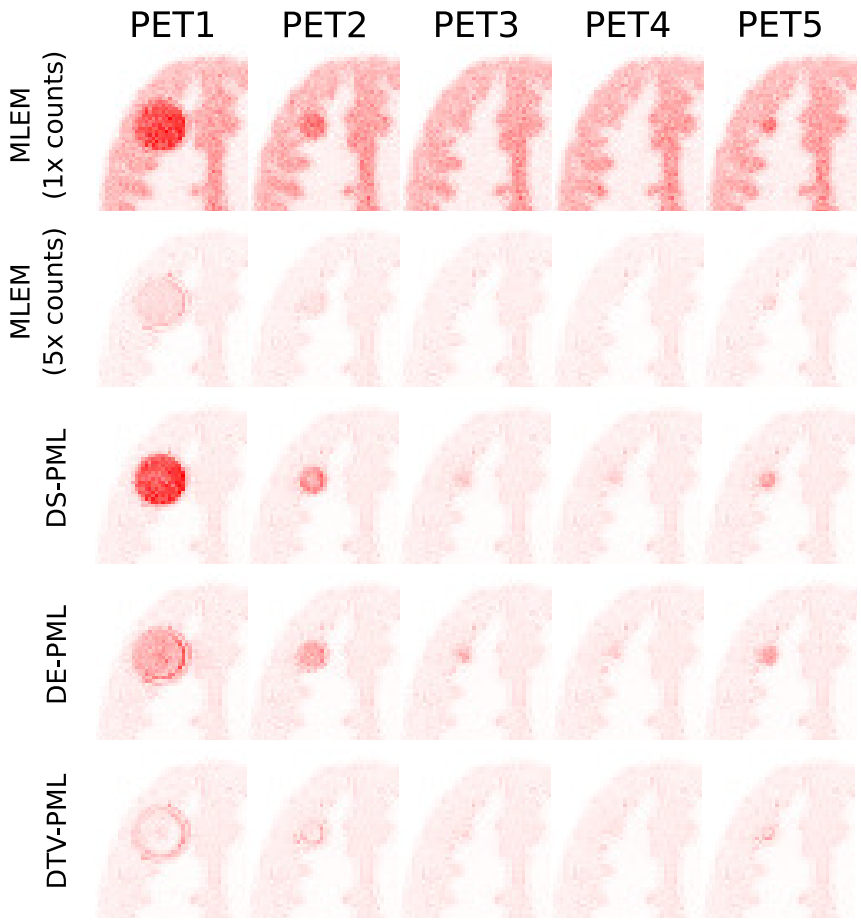


Figure 4.15: Mean-squared-error (MSE) maps around the tumour for MLEM with a counts level of 1, MLEM with 5 times the number of counts, and DS-PML, DE-PML and DTV-PML. β values for each of the difference-image priors are listed in Table 4.1. All of the difference-image prior reconstruction methods had $\sigma_w = \infty$. Note that MSE values are calculated as the voxel-wise bias squared plus the voxel-wise standard deviation squared.

errors in the tumour compared to DTV-PML. This reflects the observation that the tumour means are maintained with DE-PML at the expense of higher levels of tumour voxel variance (Figure 4.13).

4.6 Application to Real Data

4.6.1 Experimental Methods

Following the 2D simulation studies described above, the proposed priors were applied to a real longitudinal pair of [^{18}F]-FDG datasets acquired from a head and neck cancer patient using a Siemens mCT time-of-flight PET-CT scanner

(Siemens Healthcare, Erlangen, Germany). The patient exhibited a highly-uptaking tumour in the right side of the lower jaw, which reduced in size and uptake between the two scans (an interval of 85 days). The injected activity of each of the two scans was 340 MBq and 357 MBq, resulting in total prompt counts of 78.5×10^6 and 76.7×10^6 respectively in sinograms acquired in the head bed position. Note that although the data were acquired in time-of-flight format, sinograms were rebinned to non-time-of-flight for this experiment.

As well as the prompts emission sinograms, CT images were acquired for attenuation correction, and delayeds emission sinograms were acquired for randoms estimation. The vendor-supplied e7 Tools software (Siemens Healthcare, TN, USA) was used to calculate attenuation correction factors and randoms sinograms to be used in the reconstructions. Scatter correction was performed using single scatter simulation, as in [236].

In order to use the original datasets as double-counts datasets, counts-reduced versions of the data were produced by randomly removing individual counts with a probability of 0.5, similarly to the counts-reduction method described in Section 3.3.1. The randoms and scatters estimates were scaled by a factor of 0.5 accordingly.

The counts-reduced longitudinal datasets were then reconstructed with DS-PML, DE-PML, and DTV-PML, using projectors based on the mCT scanner geometry [236]. To reduce reconstruction time, these reconstruction methods were accelerated with an ordered subsets (OS) implementation [237]. The same counts-reduced datasets were also reconstructed with the OS version of MLEM (OSEM). All reconstructions were run with 5 iterations of 21 subsets reconstructing into an image grid size of $400 \times 400 \times 109$, with voxel sizes of $2.036 \times 2.036 \times 2.027\,\text{mm}^3$. Because of the large degree of noise present in head and neck PET images (due to the relatively low metabolic activity of tissues in the neck), these reconstructions included a small amount of resolution modelling according to Equation 2.10, with the PSF approximated by an isotropic Gaussian kernel of 3 mm FWHM (compared to the 4.1 mm FHWM previously measured as the resolution of the mCT scanner [238]).

For the difference-image prior methods the weights w_{sk} were 1 for all s

and k, and the total activity correction factors, α_s, were given by T_1/T_s where T_s denotes the total number of prompt counts recorded in scan s. For DS-PML, β was varied between 0 and 0.2 and ε was 1×10^{-6}. For DE-PML, β was varied between 0 and 2000 with a Gaussian Parzen window width W of 0.02. Finally, for DTV-PML, β values were varied between 0 and 0.1, and $\varepsilon = 1 \times 10^{-6}$. Whereas in the 2D simulation studies only DE-PML used a mask for the calculation of the penalty gradient, masking was required for all priors for the real 3D datasets due to the low sensitivity at the edges of the field-of-view. This penalty mask was calculated using the spatial extent of the patient's anatomy based on the PET1 attenuation map, and cropped to remove the low-sensitivity axial edges of the field-of-view.

To estimate the alignment operators required in Equations 4.2 and 4.4, non-rigid image registration of the two CT-based attenuation maps was performed utilising a demons-based registration (MATLAB R2014b, MathWorks, MA, USA). The PET1 image space was selected as the reference space so that $\boldsymbol{M}_1$ was equal to the identity matrix. The output of the demons-based registration of the PET2 CT image to the PET1 CT image was a displacement field, which was converted into matrix form using trilinear interpolation to obtain $\boldsymbol{M}_2$.

The quality of the reconstructed images was assessed by measuring the mean of the tumour and the CV in a background region in both the PET1 and PET2 reconstructed images. To serve as a reference, the original full-counts data were reconstructed using OSEM with 5 iterations of 21 subsets, and the tumour mean and background CV were measured.

4.6.2 Results

The tumour mean and background CV for the real data experiment are shown in Figure 4.16. Similarly to the results seen in the simulation experiments, increasing β reduces background noise while affecting the tumour mean. In the PET1 reconstructions, DS-PML and DTV-PML introduced a negative bias into the tumour mean, whereas DE-PML produced a positive bias. In the PET2 reconstructions, DS-PML and DTV-PML produced tumour mean

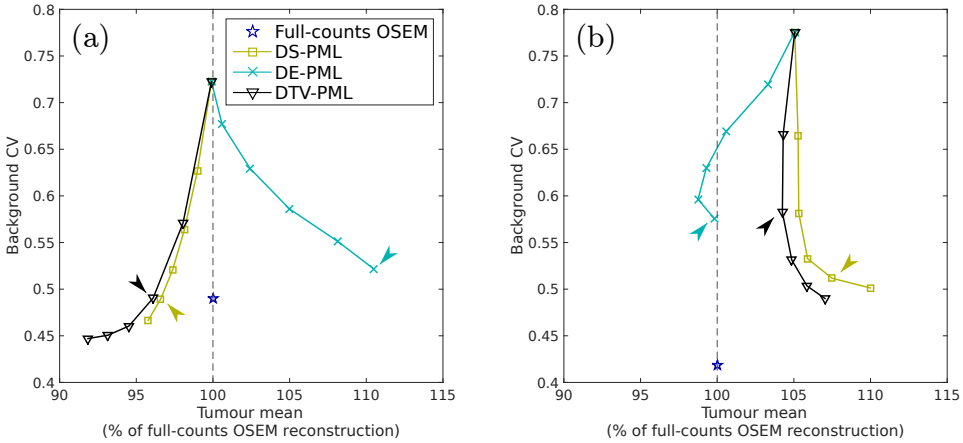


Figure 4.16: Tumour mean vs background CV as a function of β for DS-PML (yellow squares), DE-PML (cyan crosses) and DTV-PML (black triangles) for the real data experiment. Tumour mean values are shown as a percentage of the double-counts OSEM reconstruction (blue stars). In the PET1 reconstructions (a), when using the difference-image prior methods the CV in the background was reduced to levels similar to the double-counts reconstruction, however in PET2, the background CV did not reduce so greatly. Arrows for each method indicate the β values used in the reconstructions shown in Figures 4.17–4.18.

values similar to unpenalised OSEM (i.e. with $\beta = 0$), which was itself positively biased compared to the double-counts OSEM reconstruction. It was also observed that in general the PET2 background CV did not fall by the same amount as the PET1 background CV.

Based on Figure 4.16, β values were selected for the real data case. These values were 0.16 for DS-PML, 2000 for DE-PML, and 0.04 for DTV-PML. Figure 4.17 shows the reconstructed images for each of these methods, alongside the double-counts OSEM reconstruction and the OSEM reconstruction with the same data as the difference-image prior methods. It was noted that the tumour in the PET2 image consisted of two distinct peaks as opposed to the single peak observed in the PET1 images. The background noise reduction indicated by Figure 4.16 is visible in the difference-image prior methods, particularly in the PET1 scans. The appearance of the tumours generally appears the same as the OSEM methods, although DTV-PML smooths the tumour slightly. The distinction between the three methods is most pronounced when considering the difference images though, where the mechanism of each of the three priors becomes apparent. The DS-PML method successfully sets a large

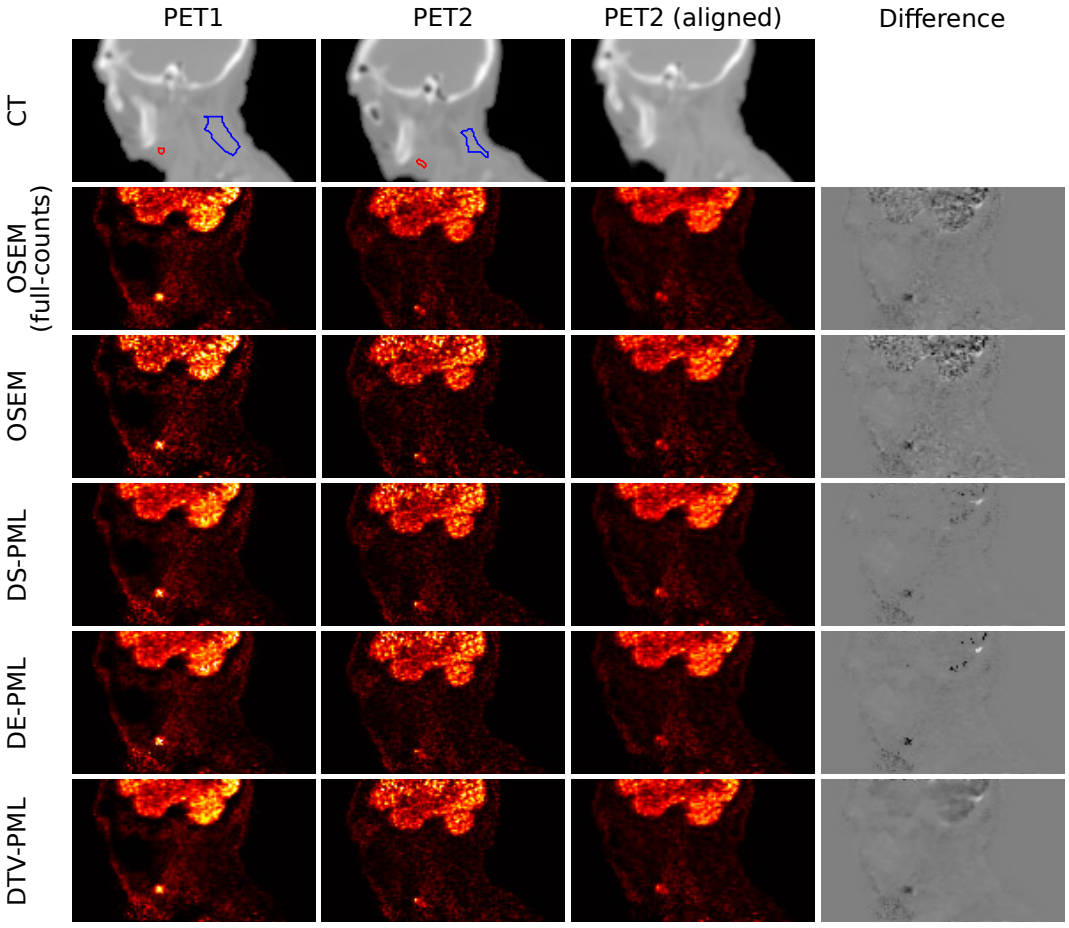


Figure 4.17: Example images for the real data case study. From left to right: PET1 images in their native space; PET2 images in their native space; PET2 images aligned to the PET1 image space; and difference images in the PET1 image space (i.e. column 3 minus column 1). Top row shows CT images with tumour (red) and background (blue) regions of interest shown for reference. PET reconstructions include the double-counts OSEM reconstructions, OSEM reconstructions, and reconstructions using DS-PML, DE-PML and DTV-PML, each with β values as indicated in Figure 4.16. These latter three difference-image prior methods reduce noise in the background of the image while maintaining the visual appearance of the tumour. Each also displays a distinctive difference image reflecting the characteristics encouraged by each prior.

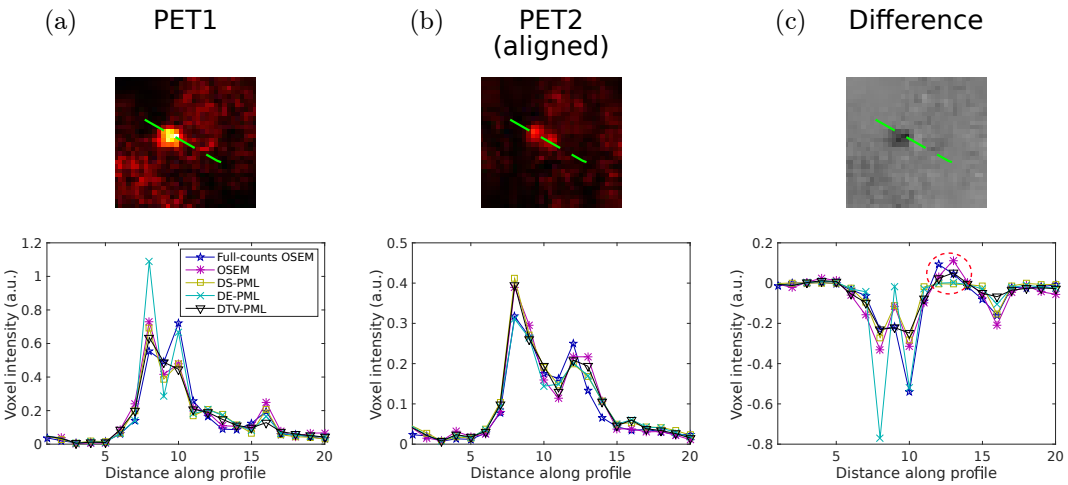


Figure 4.18: Line profiles through the tumour for (a) the PET1 images, (b) the aligned PET2 images, and (c) the difference images (bottom). The top row shows profile locations in the double-counts OSEM reconstructions. The dashed red circle highlights differences due to a region that increased in uptake from PET1 to PET2 that was suppressed by the DS-PML and DE-PML methods but retained by the DTV-PML method.

portion of the difference image to zero (or close to zero), enhancing the visibility of the remaining changes. DE-PML has also set much of the difference image background to zero, with relatively few remaining intense voxels, mainly in the tumour. Lastly, the DTV-PML method has effectively smoothed the difference image. It is noted that in this real data case, there remained other large changes in the difference image, most noticeably around the cerebellum.

Finally, Figure 4.18 shows profiles through the tumour in the PET1, aligned PET2, and difference images for each reconstruction method. In the PET1 DE-PML reconstruction (Figure 4.18(a)), the tumour peak is considerably higher than the other reconstruction methods, reflecting the positive bias observed in Figure 4.16. In the PET2 reconstructions (Figure 4.18(b)), the overestimation of the tumour mean by OSEM, DS-PML and DTV-PML is reflected in the left peak of the tumour profile. The difference image profiles (Figure 4.18(c)) show the overestimation of the difference provided by DE-PML, and also the effect of each method in the smaller peak of the PET2 tumour, which was more intense in PET2 than in PET1 (red circle). DS-PML and DE-PML suppress this small real change relative to the OSEM reconstructions, whereas DTV-PML better preserves it.

4.7 Discussion

This Chapter has built upon the concept of simultaneous reconstruction of longitudinal PET datasets via the use of difference-image penalty terms introduced in Chapter 3. Specifically, the methods have been extended to accommodate an arbitrary number of scans, account for varying scan activity levels and misalignment, and allow the use of novel penalties that encourage low entropy difference images (DE-PML) and difference images with sparse spatial gradients (DTV-PML). The results of 2D simulation study experiments showed that using these longitudinal difference-image priors to reconstruct S longitudinal scans can reduce image noise up to levels typically achieved by standard reconstructions using $S\times$ the number of recorded counts. However, there is a tradeoff between the level of background noise reduction achieved and the reconstruction of regions of change between scans.

In particular, the results showed that all of the proposed penalty terms introduce some level of bias in terms of estimation of tumour mean values with sufficiently high penalty strengths (Figure 4.4). Nonetheless, the entropy-based method DE-PML was observed to produce the lowest levels of bias when using a high-counts MLEM reconstruction as the reference. On the other hand, when considering reconstruction error in and around the tumour, the DTV-PML method produced the lowest %RMSE values, outperforming even the DE-PML method (Figure 4.5). Given that the DE-PML method reconstructs tumour means more accurately, the improvement in %RMSE seen with the DTV-PML method could be inferred to result from noise reduction within the tumour, an interpretation reinforced by Figures 4.6 and 4.9. Overall though, the level of bias introduced by any of the proposed methods was small, and inspection of the longitudinal trends of tumour mean (Figure 4.8) showed that for the particular case used (intensity reduction of 25% and radius reduction of 47%), even a relatively strong regularisation in terms of high levels of background noise reduction allows for clear discernment of the longitudinal behaviour of the tumour.

When the regularisation level was held constant and tumour response varied, the performance of the proposed methods was largely stable. At large radius changes (with fixed intensities), error levels began to rise, with DTV-PML producing the lowest errors of the three penalties (Figure 4.10). This is due to the noise reduction properties discussed previously. In general, with any of the three methods, error levels in the tumour as a function of response are more complex than for MLEM reconstructions, with the reconstruction error in each scan's tumour depending on the characteristics of the tumour in other scans. Before use in a clinical setting, the proposed methods would need to be validated across a representative, application-specific set of tumour responses to ensure that they are robust to the range of changes anticipated in that context. Such a validation study was beyond the scope of this thesis.

When the proposed difference-image prior methods were extended to the five-scan 2D case results were consistent with the results seen previously with the two-scan case. In particular, the tradeoff between background noise reduction and tumour bias was shown to apply for varying σ_w, which controlled the coupling between scans throughout the longitudinal series (Figure 4.12). At $\sigma_w = \infty$, maximum noise reduction was achieved at the expense of bias levels similar to those observed in the two-scan case. Also in agreement with the two-scan case was the ability of the DE-PML method to more accurately reconstruct tumour means (Figure 4.12), and the error reduction of DTV-PML due to noise reduction in the tumour (Figure 4.15). In the five-scan case, the visual impact of using difference-image priors is stronger in the reconstructed images due to the greater amount of data that is available to be shared between the scans.

The results of applying the proposed methods to a real data case study generally reflected those of the simulation experiments, with increasing β reducing background image noise and preserving tumour appearance. In this case though, the degree of noise reduction was not equivalent in each scan, with the PET1 noise reduced by a greater amount than the PET2 when using the proposed methods (Figure 4.16). The reason for this may be in the asymmetrical nature of the alignment operators; by calculating the penalty

gradient in the PET1 space, the effects of interpolation were avoided. On the other hand, for the PET2 image, the one-step-late penalty gradient term added to the sensitivity image would be slightly blurred by interpolation, reducing the effect of the penalty. Future experiments could investigate this effect by reconstructing the images with the alignment reversed, i.e. using PET2 as the reference space, or even experiment with the use of both or intermediate spaces.

In terms of tumour quantification, some bias was introduced into the mean tumour values relative to the standard OSEM reconstructions (Figure 4.16), similarly to the 2D simulation study results. However, the real data experiment also highlighted some of the potential issues with the three difference-image priors, including in the case where there are additional changes between the images beyond those caused by the tumour response, such as those that were observed in the brain of the patient (Figure 4.17).

When using the DS-PML method, the valid differences caused by the appearance of a second tumour peak in the PET2 scan were seen to be suppressed (Figure 4.18(c)). This could be due to the small amplitude of this peak, relative to the range of changes due to noise present elsewhere in the image. The effect of this would be that a larger β would be required to perform adequate noise reduction in the difference image, simultaneously suppressing the small true change in this tumour peak. A similar effect was seen for the DE-PML method, where the small peak was suppressed in the difference image to bring these voxels into the background peak in the difference histogram. In addition, for the DE-PML method the existence of large changes elsewhere in the image, particularly in the brain, is hypothesised to be the cause of the large observed biases in tumour mean measurements (Figure 4.16). By attempting to produce a low entropy difference image, similar changes can be encouraged to have the same value, even over large spatial distances. This is reflected in Figure 4.17, where it can be seen that the changes in the brain and the tumour in the DE-PML difference image seem to have similar values, despite them appearing distinct in the other difference images. Contrary to the other two methods, the DTV-PML method seems to deal better with the existence

of both higher noise and other changes in the difference image, maintaining the second tumour peak that was suppressed by the other two methods (Figure 4.18(c)). However, the DTV-PML method still introduced bias into the tumour reconstructions because of spatial smoothing applied by the penalty in the difference image domain. This reflects the results seen in the simulation studies (e.g. Figure 4.6), where the tumour was smoothed relative to the standard MLEM reconstructions. In the real data reconstructions, the small size of the tumours could have augmented this effect.

The introduction of bias into tumour mean values is certainly a limitation of the methods as presented in this Chapter. However, as shown in this work, these biases are relatively small relative to typical tumour variations due to disease progression or tumour response. Furthermore, the amount of bias that is acceptable will be application specific. For example, in tumour detection contexts, the benefit of background noise reduction provided by the proposed methods could outweigh the detriment of the bias introduced into the tumour. On the other hand, in tumour response scans where the threshold for responder or non-responder can be very fine these methods might be of limited benefit (or at least lower penalty strengths would need to be used).

4.8 Summary

This Chapter presented an extension of the methods proposed in Chapter 3. In particular, an updated theoretical framework was presented which a) allowed a general number of scans to be reconstructed simultaneously, b) accounted for the effects of varying image activity and misalignment between scans, and c) allowed the penalisation of any desired characteristic of differences to be used. We proposed two new such penalties, encouraging low entropy difference images (the DE prior), and difference images with low total variation (the DTV prior). These priors were tested alongside the ℓ_1-norm sparsity-encouraging prior proposed in Chapter 3 (the DS prior). The results of simulation studies showed that considerable noise reduction can be achieved across background portions of the image by using these methods. The level of this noise reduc-

tion is related to the number of scans reconstructed, with S equal-dose scans providing a maximum noise reduction similar to increasing counts levels by S times when using standard, independent-dataset reconstruction methods. The observed effects of the reconstruction methods varied within a longitudinally changing tumour, with the DE prior maintaining low levels of bias, while the DTV prior provides the lowest reconstruction errors by way of noise reduction at the cost of additional tumour bias. Tumour reconstruction error with the proposed methods was observed to be relatively stable over a range of tumour behaviours. To demonstrate the feasibility of using the proposed methods clinically, a two-scan longitudinal head and neck cancer patient data-series was reconstructed with the proposed methods, including the relevant corrections for varying total activity levels and misalignment. In agreement with the simulation study results, background noise was reduced throughout the images when using the proposed methods, at the cost of induced tumour biases of varying degrees. For the DE prior this bias was large, and is hypothesised to be due to the global nature of the entropy, which allows distant large changes to affect the reconstruction of tumour regions. In addition, both the DS and DE priors suppressed a low amplitude peak in the tumour difference image, due to the relative amplitude of this peak compared to image-wide noise. Conversely, the DTV prior provided a good reconstruction of the tumour, preserving this small peak.

Future work includes the continued development and investigation of alternative priors to act on difference images. In addition, further investigation is required into the effect of using various image registration methods on the PET image reconstruction accuracy, since the minimisation of errors in the alignment operators is expected to be important for the proposed methods to work optimally. This would have to be performed in an application-specific manner due to the range of application-specific registration models available in the literature [239]. In addition, this application-specific testing using large cohorts of patients would be important before clinical usage in order to ascertain which applications could benefit from the proposed methods.

Chapter 5:
Guided Multi-Tracer PET Image Reconstruction

Chapters 3 and 4 proposed methods of improving longitudinal PET images via simultaneous reconstruction techniques that penalised certain characteristics of the difference images. As discussed in Section 2.4.1, guided PET image reconstruction is an alternative approach for sharing information between datasets, where information is extracted from a fixed prior image and embedded in the reconstruction of the target PET image. Traditionally, guided PET image reconstruction has used anatomical images from MR or CT scans to provide this guidance information. However, anatomical guidance of PET image reconstruction may not always be possible or optimal due to a number of reasons, such as a lack of sufficiently high contrast MR imaging in a study protocol and/or the poor quality of low-energy CT images acquired for PET attenuation correction purposes only.

In multi-tracer PET studies as described in Section 1.2.2, there is the opportunity to transfer information between radiotracer distributions in a manner similar to anatomically guided PET image reconstruction. In particular, many multi-tracer PET protocols include [^{18}F]-FDG scans which are non-specific in their uptake, offering a level of structure and contrast that is expected to permit guided image reconstruction of other radiotracer images.

In this Chapter we demonstrate guided reconstruction for multi-tracer PET, investigating the specific dual-tracer case of [^{18}F]-FDG/[^{11}C]-methionine imaging of brain tumours [17, 98]. Often the reconstructed [^{11}C]-methionine

images are very noisy and require large amounts of post-reconstruction smoothing to obtain clinically useful images (Figure 5.1(b)). On the other hand, the [^{18}F]-FDG images generally exhibit high contrast and low noise (Figure 5.1(a)). This Chapter investigates the use of the higher quality [^{18}F]-FDG images to guide the reconstruction of the lower quality [^{11}C]-methionine data, referred to here as FDG-guided [^{11}C]-methionine image reconstruction. To achieve this guidance, a weighted quadratically penalised MAP objective function was used, where the weights are derived from a pre-existing [^{18}F]-FDG image, under the assumption that [^{18}F]-FDG and [^{11}C]-methionine distributions are structurally similar, i.e. that edges in [^{18}F]-FDG images correspond to those in [^{11}C]-methionine images (Figure 5.1). This assumption is supported by the observation that both of these non-specific radiotracers tend to conform to anatomical boundaries, with increased uptake in cortical tissues [17]. A 3D simulation study was performed to explore the effects of the various hyperparameters of the proposed reconstruction method, and preliminary tests were performed to validate the simulation results on real patient data. This Chapter is based on work published by the author during the course of their PhD studies [240].

5.1 FDG-Guided [^{11}C]-Methionine Image Reconstruction

In this Chapter, one of the simplest guided reconstruction approaches is utilised in order to demonstrate the principle of multi-tracer guided image reconstruction. Namely, the MAP objective function (Equation 2.22) with the weighted quadratic penalty (Equation 2.40) is used.

Explicitly, the objective function is

$$\Phi^{\mathrm{MAP}}(\boldsymbol{\theta}) = \Phi^{\mathrm{ML}}(\boldsymbol{\theta}|\boldsymbol{m}) - \beta U(\boldsymbol{\theta}|\boldsymbol{\phi}), \tag{5.1}$$

where $\boldsymbol{\theta}$ is here the [^{11}C]-methionine image and $\boldsymbol{\phi}$ is the guiding [^{18}F]-FDG

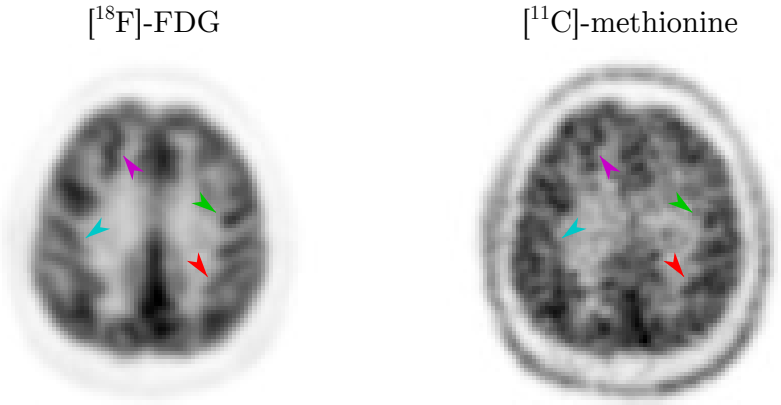


Figure 5.1: Example [^{18}F]-FDG/[^{11}C]-methionine brain images, reconstructed with 100 iterations of MLEM followed by a 5 mm FWHM Gaussian smooth to obtain clinically typical image appearance. Although the [^{11}C]-methionine image is considerably noisier than the [^{18}F]-FDG image, with a lower level of contrast, both images exhibit radiotracer uptake that conforms to grey matter/white matter boundaries, with specific structural features indicated by coloured arrows.

image, included in the penalty for clarity. The guided weighted penalty $U(\boldsymbol{\theta}|\boldsymbol{\phi})$ is then given by:

$$U(\boldsymbol{\theta}|\boldsymbol{\phi}) = \frac{1}{4}\sum_{j=1}^{J}\sum_{k\in\mathcal{N}_j} w_{jk}(\boldsymbol{\phi})\,(\theta_j - \theta_k)^2\,, \tag{5.2}$$

where the weights are given explicitly as a function of the [^{18}F]-FDG image $\boldsymbol{\phi}$. By calculating w_{jk} from the [^{18}F]-FDG image, information about which voxels are *a priori* similar can be built into the [^{11}C]-methionine image reconstruction.

In this Chapter, the prior weights are calculated using Gaussian kernel functions, similar to those used in the kernelised EM method [181–186] and given in Equation 2.42. In terms of the [^{18}F]-FDG image $\boldsymbol{\phi}$, the weights are given by

$$w_{jk}(\boldsymbol{\phi}) = \exp\left(\frac{-\|\boldsymbol{f}_j(\boldsymbol{\phi}) - \boldsymbol{f}_k(\boldsymbol{\phi})\|_2^2}{2\sigma_f^2}\right)\exp\left(\frac{-\|\boldsymbol{r}_j - \boldsymbol{r}_k\|_2^2}{2\sigma_s^2}\right), \tag{5.3}$$

where $\boldsymbol{f}_j(\boldsymbol{\phi})$ is the feature vector for voxel j based on image $\boldsymbol{\phi}$ and $\boldsymbol{r}_j$ is the spatial location of voxel j, with corresponding definitions for $\boldsymbol{f}_k(\boldsymbol{\phi})$ and $\boldsymbol{r}_k$. The hyperparameters σ_f and σ_s control the widths of the two kernels. The

first kernel is the same as that in Equation 2.42, based on feature vectors, and the second is a spatially modulating kernel, which can be removed by setting $\sigma_s = \infty$.

In this Chapter each feature vector, $\boldsymbol{f}_j(\boldsymbol{\phi})$, contains the intensities from the $p \times p \times p$ patch centred at voxel j, arranged as a column vector. Importantly, the feature vectors in this Chapter are normalised by the standard deviation of each element across the image, as suggested by Wang and Qi [180], so that:

$$f_{j,m} = \frac{\hat{f}_{j,m}}{\mathrm{SD}_m(\{\hat{\boldsymbol{f}}_j\}_{j=1}^J)}, \tag{5.4}$$

where $\hat{\boldsymbol{f}}_j$ are the unnormalised feature vectors, and $\mathrm{SD}_m(\{\hat{\boldsymbol{f}}_j\}_{j=1}^J)$ is the standard deviation of the mth element of the unnormalised feature vectors over all voxels. This normalisation is useful for two reasons, as demonstrated in Figure 5.2. Firstly, normalisation ensures that the feature space is isotropic by rescaling each element relative to its spread, allowing equal contribution to the Euclidean distance from the different elements of the feature vectors (Figures 5.2(a) and (c)). Secondly, normalisation results in unitless Euclidean distances, $\|\boldsymbol{f}_j - \boldsymbol{f}_k\|_2$, which allows σ_f to be defined without consideration of the scale of the feature vector elements (Figures 5.2(b) and (d)). For PET image patches, the scale of the feature vector elements could be normalised by patient weight and injected activity, i.e. by using SUVs, but in this Chapter we use the standard deviation approach described above to retain a consistent framework where any imaging information could be used in a corresponding manner. Note that by using image patches as features in this Chapter, the isotropy of feature space is not a consideration since all elements of the feature vectors automatically have the same scale. Also note that in this Chapter, positions are not normalised so that σ_s is expressed in the same units as the spatial distance between $\boldsymbol{r}_j$ and $\boldsymbol{r}_k$.

In the kernel method discussed previously in Section 2.4.1, sparsification of the kernel matrix is performed in order to reduce computational costs and to reduce the effect of erroneous similarities [180]. In the original proposal of the method, suggested sparsification methods included the use of image neighbour-

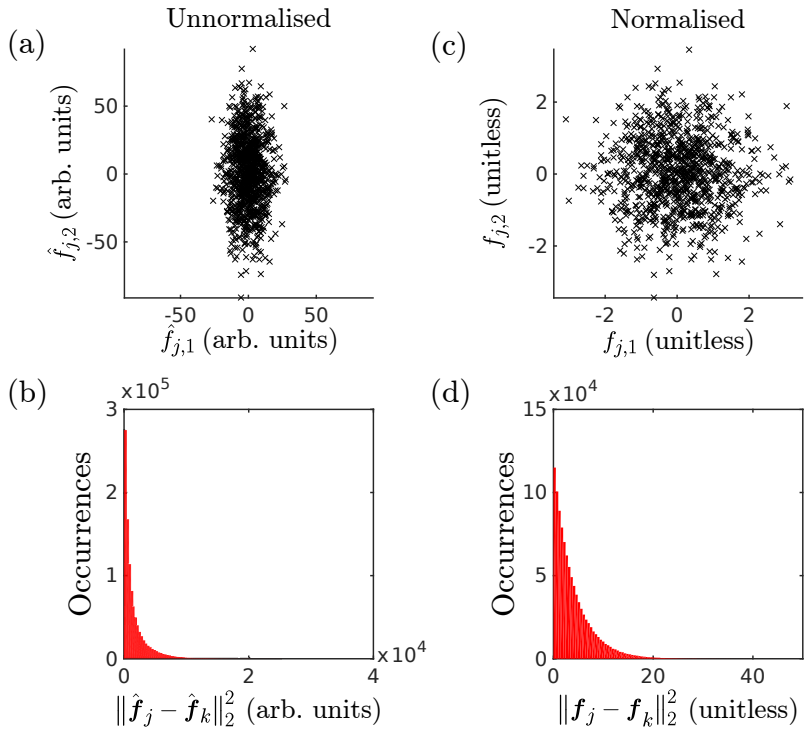
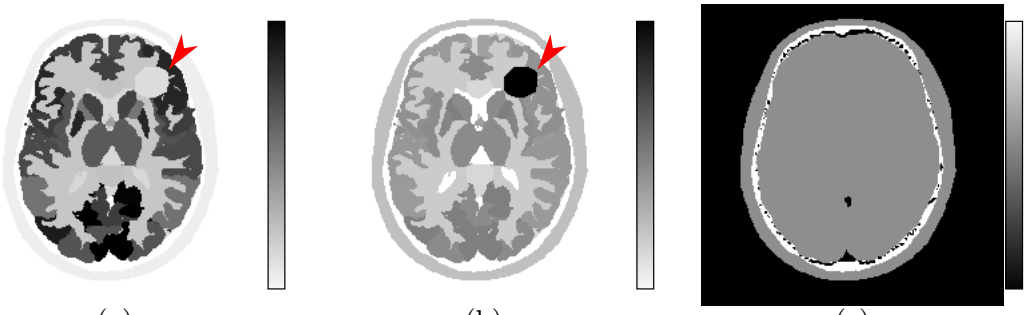


Figure 5.2: The effects of feature vector normalisation. For demonstration purposes, the unnormalised two-element feature vectors, $\hat{\boldsymbol{f}}_j = [\hat{f}_{j,1}, \hat{f}_{j,2}]^T$, are random draws from Gaussian distributions of differing standard deviations. (a) If the scale of $\hat{f}_{j,1}$ is larger than that of $\hat{f}_{j,2}$, then $\hat{f}_{j,1}$ will contribute proportionally more to the Euclidean distance in Equation 5.3. (b) Furthermore, for unnormalised feature vectors, the Euclidean distance between pairs has arbitrary scaling and units, which must be accounted for when selecting σ_f values. (c) By normalising feature vectors (Equation 5.4), the feature space becomes more isotropic, allowing equal weighting from different feature vector elements. (d) In addition, the range of differences becomes normalised and unitless, allowing selection of σ_f without explicit consideration of the scale or units of features being used.

hoods, k-nearest neighbours or the threshold-based ε-ball method (Equation 2.44). Similarly, in the Bowsher prior, sparsification is achieved by keeping only the B most similar voxels for each voxel and setting these similarities to 1, while all other weights are set to 0 [169]. In this Chapter, an implicit two-stage sparsification using neighbourhoods and maximum values was selected. The initial neighbourhood sparsification is implicit in the definition of Equation 5.2, where penalty contributions are only taken between voxels within the neighbourhood $\mathcal{N}_j$ around voxel j. The second sparsification step keeps only the n_{Max} maximum value weights for each voxel, and sets all other weights to zero, similarly to both the Bowsher prior and a recent implementation of the



(a) (b) (c)

Figure 5.3: High resolution (a) [^{18}F]-FDG and (b) [^{11}C]-methionine ground truth images for the 3D simulation study to characterise the effects of each hyperparameter listed in Table 5.1. Note the tumour in the inferior aspect of the right frontal lobe (red arrows). (c) the attenuation map used for all simulations, showing three voxel classes: air, water and bone.

kernel method [186].

To maximise the objective function in Equation 5.1 the modified EM approach of De Pierro [153] was implemented, as given previously in Equations 2.36–2.39.

5.2 3D Simulation Study

To evaluate the effect of guiding [^{11}C]-methionine PET image reconstructions with [^{18}F]-FDG images a 3D simulation study was initially performed. Guided reconstruction performance was measured as a function of the various reconstruction hyperparameters and compared to standard MLEM reconstructions followed by both Gaussian and [^{18}F]-FDG-derived non-local-means (FDG-NLM) filters (see Section 2.4.1).

Data simulation and image reconstruction were implemented in MATLAB R2016a (MathWorks, MA, USA), using in-house Siddon-based mMR projectors [228, 229].

5.2.1 Experimental Methods

Two ground truth 3D brain images were generated using the freely available NCAT brain phantom phantom [227], with the contrasts adjusted to emu-

late [^{18}F]-FDG and [^{11}C]-methionine scans. The matrix size for these ground truth images was $688 \times 688 \times 254$, with a voxel size of $1.04313 \times 1.04313 \times 1.015625\,\text{mm}^3$. A tumour was introduced into the ground truth images, extending across the border between white and grey matter in the inferior aspect of the right frontal lobe (Figure 5.3). The tumour was hypo-intense in the [^{18}F]-FDG image with an intensity of 60% that of adjacent white matter, and in the [^{11}C]-methionine image the tumour was hyper-intense with a tumour to grey matter ratio of 2:1.

Using these ground truth images, two Siemens Biograph mMR datasets were simulated, using the same data generation pipeline: First, each ground truth image was blurred by a Gaussian PSF of 4.3 mm FWHM [217] before projecting them into the standard mMR sinogram space. Attenuation sinograms were then produced using the segmented attenuation map (Figure 5.3(c)), with the following tissue/value pairs for the linear attenuation coefficient: air, $0\,\text{cm}^{-1}$; water, $0.096\,\text{cm}^{-1}$; and bone, $0.172\,\text{cm}^{-1}$ [106]. These attenuation factors were applied to the forward projected, resolution degraded sinograms, along with normalisation factors from a real mMR scan. The additive contribution from scattered events was modelled as a smoothed copy of the noise-free forward projection of the blurred images, and randoms were modelled as uniform sinograms. The counts levels and scatter and randoms fractions were adjusted so that the final noise-free [^{18}F]-FDG dataset contained 300 M counts, with 20% scatter and 30% randoms, and the final noise-free [^{11}C]-methionine dataset contained 70 M counts, with 20% scatter and 25% randoms. From these noise-free datasets, noisy datasets were produced by introducing Poisson noise.

To produce the prior image to be used to guide the reconstruction of the [^{11}C]-methionine dataset, the [^{18}F]-FDG dataset was reconstructed into the standard mMR $344 \times 344 \times 127$ image grid with a voxel size of $2.08626 \times 2.08626 \times 2.03125\ \text{mm}^3$, using 100 iterations of MLEM without resolution modelling, and then smoothed by a 4 mm FWHM Gaussian kernel. This [^{18}F]-FDG prior image was used to calculate the weights according to Equation 5.3 in order to guide the reconstruction of the [^{11}C]-methionine images into

the same image grid, (also without resolution modelling). Note that to improve computational efficiency and to avoid including air voxels in the feature normalisation (Equation 5.4), features were extracted and weights calculated only within an image mask based on the attenuation map. For this simulation study, to reduce the large hyperparameter space of the reconstructions, all feature vectors were $3 \times 3 \times 3$ cubic patches arranged as 27×1 vectors, and all neighbourhoods, $\mathcal{N}_j$, were $5 \times 5 \times 5$. Therefore the guided [^{11}C]-methionine reconstructions were a function of β, n_{Max}, σ_f, and σ_s. To explore the hyperparameter space, a baseline value for each of these variables was defined (see Table 5.1), and each one was then varied between a minimum and maximum value in turn. These FDG-guided MAPEM reconstructions were run to 200 iterations of Equation 2.36 without resolution modelling, initialised with one iteration of standard MLEM to avoid convergence artefacts at the axial edges of the field-of-view.

Reconstruction error was quantified by using the regional percentage mean absolute error (%MAE) metric:

$$\%\text{MAE}_\Omega = \frac{100\%}{N_\Omega} \sum_{j \in \Omega} \frac{\left|\theta_j - \theta_j^{\text{Ref}}\right|}{\theta_j^{\text{Ref}}}, \tag{5.5}$$

where Ω is the region in which error is to be measured, N_Ω is the number of voxels in that region, and $\boldsymbol{\theta}^{\text{Ref}}$ is the reference image against which error is being calculated. Note that %MAE is equivalent to the %RMSE metric used in Chapters 3 and 4 when only a single noisy realisation is used. For

Hyperparameter	Baseline value	Range (min–max)
β	15000	5000–25000
n_{Max}	15	5–125
σ_f	4	0–8
σ_s	4 mm	0–8 mm

Table 5.1: Hyperparameters used for the reconstruction of simulated [^{11}C]-methionine datasets in the 3D hyperparameter characterisation study. Each hyperparameter had a baseline value which was held constant while each of the other hyperparameters were varied in turn with the range shown in the final column.

the 3D simulation study, %MAE was calculated in [^{11}C]-methionine images in the whole-brain and in the tumour alone, in both cases using a linearly interpolated downsampled version of the ground truth image as $\boldsymbol{\theta}^{\text{Ref}}$. Since this downsampling is by an exact factor of 2, this is equivalent to summing the 4 contributing voxels' intensities in the original ground truth image to obtain each voxel's intensity in the downsampled reference image (with a scale factor of 4). In addition to calculating %MAE values, the contrast recovery coefficient (CRC) of the tumour was measured according to

$$\text{CRC}(\boldsymbol{\theta}) = \frac{\text{Mean}_T(\boldsymbol{\theta}) - \text{Mean}_B(\boldsymbol{\theta})}{\text{Mean}_B(\boldsymbol{\theta})}, \tag{5.6}$$

where $\text{Mean}_T(\boldsymbol{\theta})$ is the mean voxel intensity within the tumour, and $\text{Mean}_B(\boldsymbol{\theta})$ is the mean voxel intensity in a white matter background region. The standard deviation (SD) of the voxel intensities within the tumour and the white matter background region were also measured.

For comparison to the FDG-guided MAPEM [^{11}C]-methionine reconstructions, the same datasets were also reconstructed with 200 iterations of MLEM without resolution modelling and filtered post-reconstruction with Gaussian filters of varying FWHM. In addition, the MLEM images were filtered with NLM filters corresponding to the same [^{18}F]-FDG-derived weights used in the MAPEM approach (FDG-NLM). In practice, this involved the multiplication of the MLEM reconstruction image vector with the weights w_{jk} inserted into a sparse $J \times J$ matrix and appropriately row-normalised.

In order to evaluate the effect of the convergence of the guidance [^{18}F]-FDG image reconstruction on the quality of the reconstructed [^{11}C]-methionine images, guidance weights were calculated from 4 mm Gaussian smoothed [^{18}F]-FDG images reconstructed using 50, 100 and 200 iterations of MLEM. These weights were calculated using the baseline hyperparameter values listed in Table 5.1. The [^{11}C]-methionine dataset was then reconstructed with each of these three sets of weights in turn, and the resultant images compared.

Finally, to assess the performance of the proposed reconstruction method in cases of strong disagreement between the guiding [^{18}F]-FDG PET image and

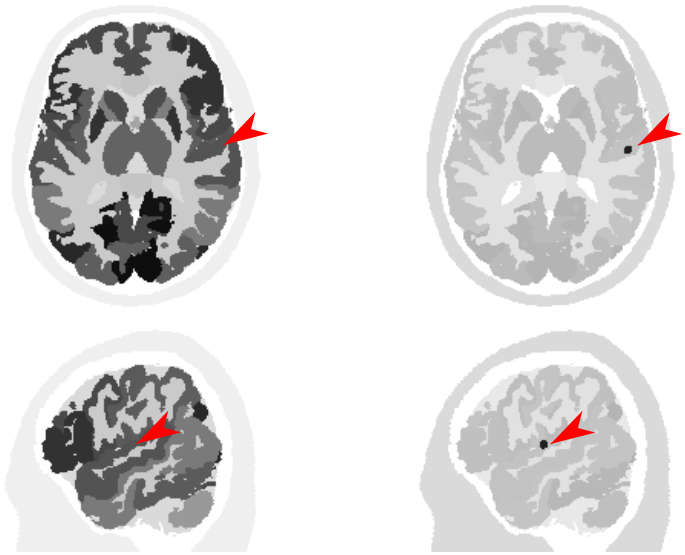


Figure 5.4: The ground truth images used for the mismatch test 3D simulation study. The [^{11}C]-methionine image (b) had a small, highly uptaking tumour (red arrows) in the grey matter, which did not match any structure in the [^{18}F]-FDG image (a). Note that attenuation maps were the same as shown in Figure 5.3 and all other aspects of the data simulation were the same as for the hyperparameter characterisation study.

the [^{11}C]-methionine image, a mismatch test was carried out using a second pair of phantoms in which there was a smaller, [^{11}C]-methionine-unique tumour embedded within the grey matter, with an intensity of 3× that of the surrounding grey matter (Figure 5.4). The corresponding region in the [^{18}F]-FDG image was set to standard grey matter values. The [^{11}C]-methionine data were reconstructed with both MLEM with a 5 mm Gaussian smooth and FDG-guided MAPEM with hyperparameters based on the hyperparameter characterisation experiments detailed above. The CRC of the [^{11}C]-methionine-unique tumour was measured for each reconstruction.

5.2.2 Results

Figure 5.5 shows the [^{18}F]-FDG prior used to calculate the weights for the hyperparameter characterisation experiments, as well as example guidance weights calculated using the baseline hyperparameter values in Table 5.1. The [^{18}F]-FDG prior exhibits good contrast between grey and white matter, as is

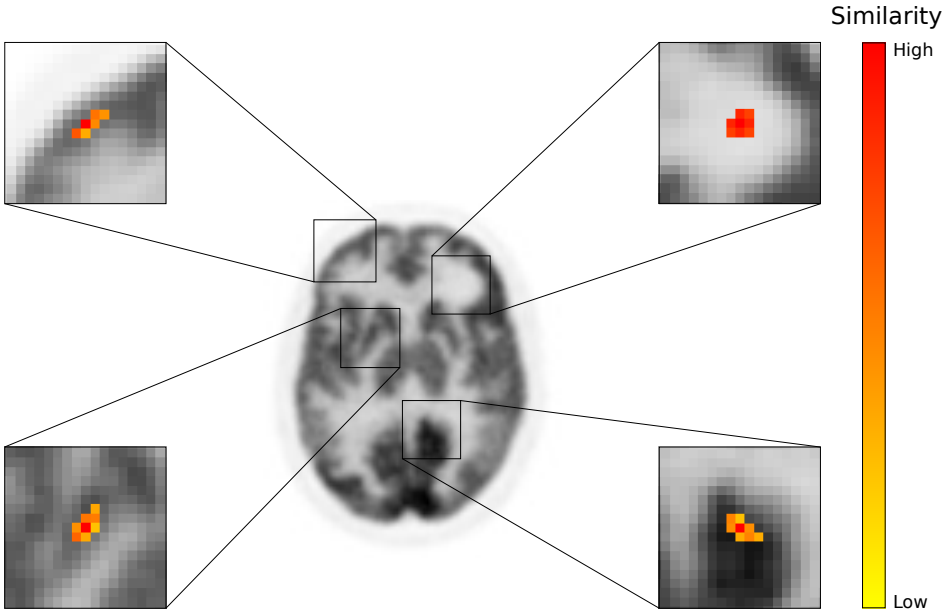


Figure 5.5: The reconstructed prior [^{18}F]-FDG image used for the 3D simulation study, and example weights w_{jk} for four choices of j, calculated for inclusion in FDG-guided [^{11}C]-methionine image reconstruction. It can be seen from the examples displayed here that the weights respect the boundaries of the grey matter and are therefore representative of the [^{18}F]-FDG image structure. In the uniform tumour region, the weights remain uniform. Note that for each example, weights that were set to zero by the sparsification step are not displayed.

typical for [^{18}F]-FDG brain images, but there is an image-wide blur due to both the PSF included in the data simulation, and the 4 mm Gaussian blur applied post-reconstruction. Nonetheless, the calculated example guidance weights reflect well the grey matter structure as intended.

Figure 5.6 shows the image metrics as a function of penalty strength, β, for the FDG-guided MAPEM. As β increases, the whole-brain %MAE in the FDG-guided MAPEM reconstructions reduces (Figure 5.6(a)), with $\beta = 10000$ providing the same whole-brain %MAE as MLEM with a 5 mm Gaussian smooth (the level of smoothing that produced the lowest whole-brain %MAE). Higher values of β reduced the reconstruction error compared to smoothed MLEM. In this case, using the FDG-NLM filter on the MLEM reconstruction produced higher whole-brain %MAE than a Gaussian filter. When considering the trade-off between tumour SD and tumour mean (Figure 5.6(b)), it is apparent that using FDG-guided MAPEM allows greater reduction in the tumour noise

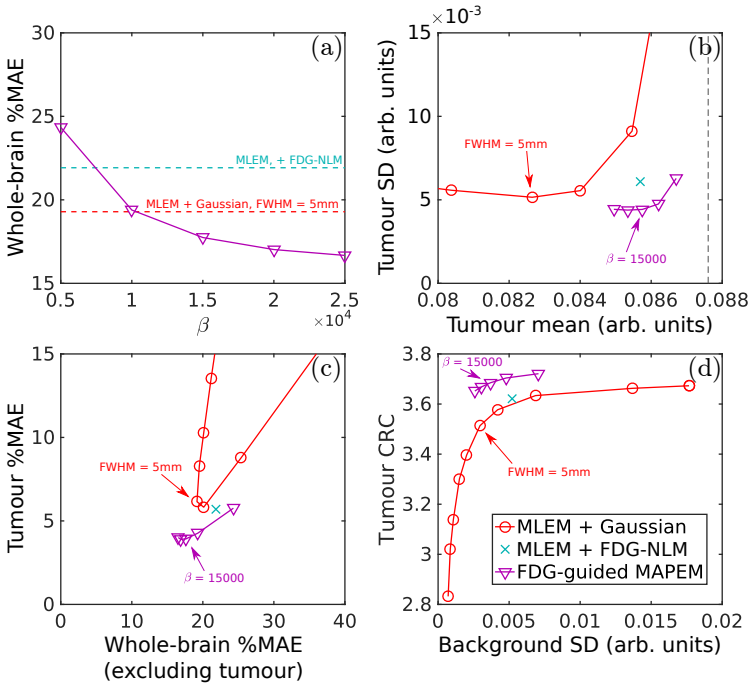


Figure 5.6: Effect of the penalty strength β on (a) the whole-brain %MAE, (b) tumour standard deviation vs tumour mean, (c) whole-brain %MAE excluding tumour vs tumour %MAE, and (d) tumour contrast recovery coefficient vs background standard deviation for FDG-guided MAPEM. Also shown are results obtained from using MLEM followed by a variable-width Gaussian smooth and MLEM followed by the FDG-NLM filter. Note that all hyperparameters (other than β) were set to the baseline values given in Table 5.1 for both FDG-guided MAPEM and the FDG-NLM filter. Dashed grey line in (b) shows the tumour mean in the ground truth image.

while maintaining the quantification of the tumour mean compared to both the FDG-NLM and Gaussian filtered MLEM reconstructions. The FDG-guided MAPEM method also outperforms MLEM with either a Gaussian or FDG-NLM filter in terms of reconstruction error in both the tumour only and the rest of the brain (Figure 5.6(c)). Specifically, comparing FDG-guided MAPEM with the baseline hyperparameter values (i.e. $\beta = 15000$ in Figure 5.6) to the 5 mm Gaussian filtered MLEM, the %MAE is reduced by 8% across the whole-brain, and 36% in the tumour. Finally, FDG-guided MAPEM produces higher tumour CRC at a given noise level when compared to the post-reconstruction filtered MLEM reconstructions (Figure 5.6(d)).

These results are reflected in the reconstructed images (Figure 5.7). The 5 mm Gaussian filtered MLEM image achieves good levels of noise reduction,

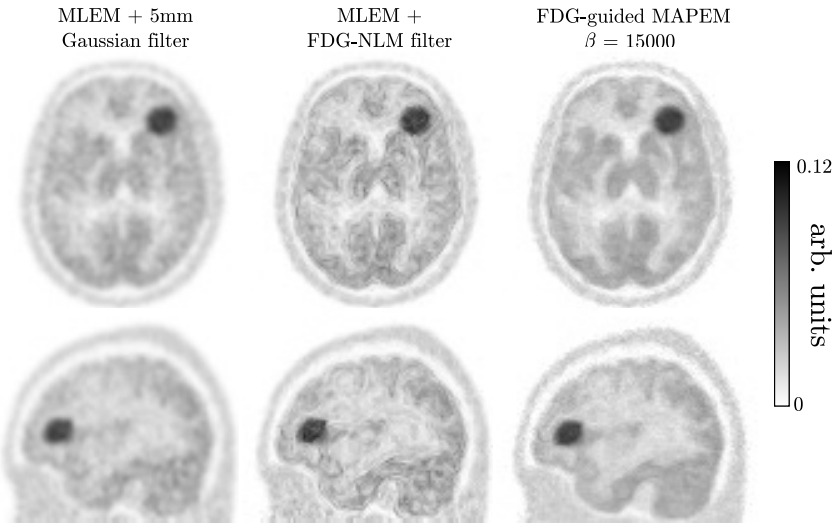


Figure 5.7: Example reconstructed images for (a-b) MLEM followed by (a) a 5 mm Gaussian filter, (b) an FDG-NLM filter, and (c) FDG-guided MAPEM with $\beta = 15000$. All undisplayed hyperparameters for the latter two methods took baseline values from Table 5.1. The FDG-guided MAPEM outperforms the Gaussian filtered MLEM in terms of noise reduction and edge-preservation while avoiding the artefacts observed in the FDG-NLM filtered MLEM image.

but at the cost of blurred edges, including around the border of the tumour. The FDG-NLM filtered MLEM image suffers less from edge degradation, but contains an artefact-like noise pattern which makes the image less visually appealing than the Gaussian filtered MLEM. Compared to both of these images, the FDG-guided MAPEM image exhibits low noise across the image while also maintaining the appearance of edges.

When the number of maximum values retained in the weights matrix per voxel (n_{Max}) was varied, the FDG-NLM filtered MLEM results greatly improve (Figure 5.8). At higher values of n_{Max} (> 40), using the FDG-NLM filter outperformed the FDG-guided MAPEM in terms of whole-brain %MAE (Figure 5.8(a)), with both methods providing lower %MAE than the 5 mm Gaussian filtered MLEM image. When the tumour SD/mean tradeoff is considered though, it is apparent that the FDG-guided MAPEM performs better than the FDG-NLM filtered MLEM method by providing a higher tumour mean for fixed standard deviation for $n_{\text{Max}} < 20$ (Figure 5.8(b)). This is reflected in a superior tradeoff between tumour CRC and background noise (Figure 5.8(d)).

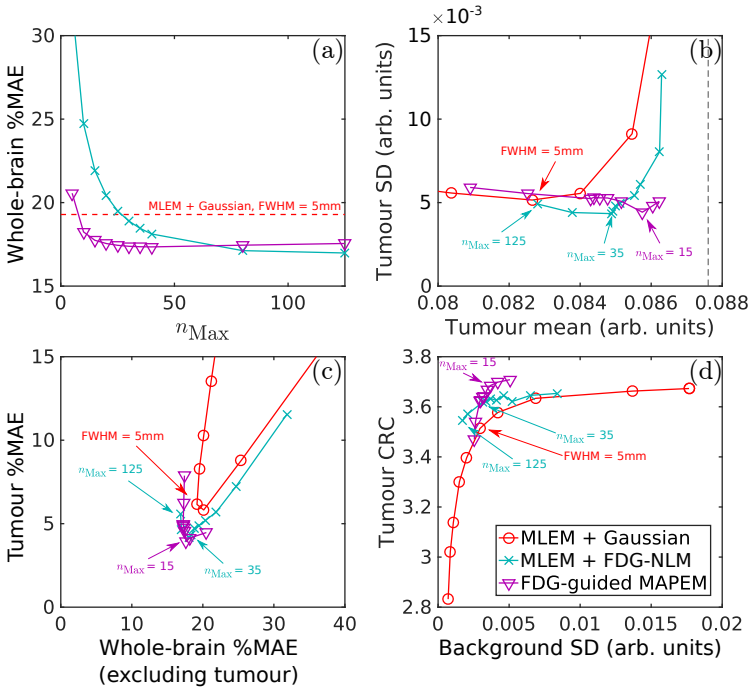


Figure 5.8: Image metrics as a function of n_{Max} for FDG-NLM filtered MLEM images and FDG-guided MAPEM. Also shown are results from Gaussian filtered MLEM. (a) Whole-brain %MAE as a function of n_{Max} for FDG-NLM and FDG-guided MAPEM, with 5 mm Gaussian filtered MLEM shown as a reference. (b) Tradeoff between tumour SD and tumour mean as a function of n_{Max} for FDG-NLM filtered MLEM and FDG-guided MAPEM, and as a function of FWHM for Gaussian smoothed MLEM. (c) Tradeoff between tumour %MAE and the %MAE in the rest of the brain as a function of n_{Max} and FWHM for the competing methods. (d) Tradeoff between tumour CRC and background noise. © 2018 IEEE

In terms of total tumour %MAE against the %MAE in the rest of the brain (Figure 5.8(c)), the FDG-NLM filtered MLEM and the FDG-guided MAPEM perform similarly, with slightly lower errors from FDG-guided MAPEM.

When considering the reconstructed images (Figure 5.9), the FDG-guided MAPEM is superior to the noise-matched FDG-NLM filtered MLEM with $n_{\text{Max}} = 35$, which retains some artefacts similar to those seen in Figure 5.7. The FDG-NLM filtered image with $n_{\text{Max}} = 125$ produces an image which is less noisy than the FDG-guided MAPEM image, but visually is less appealing due to oversmoothing and some remaining artefacts (indicated by arrows).

Figure 5.10 shows the results of varying σ_f and σ_s. Beyond $\sigma_f = 2$ and $\sigma_s = 2$ mm, the FDG-guided MAPEM results are stable, outperforming MLEM with a 5 mm Gaussian filter and a FDG-NLM filter with any corresponding

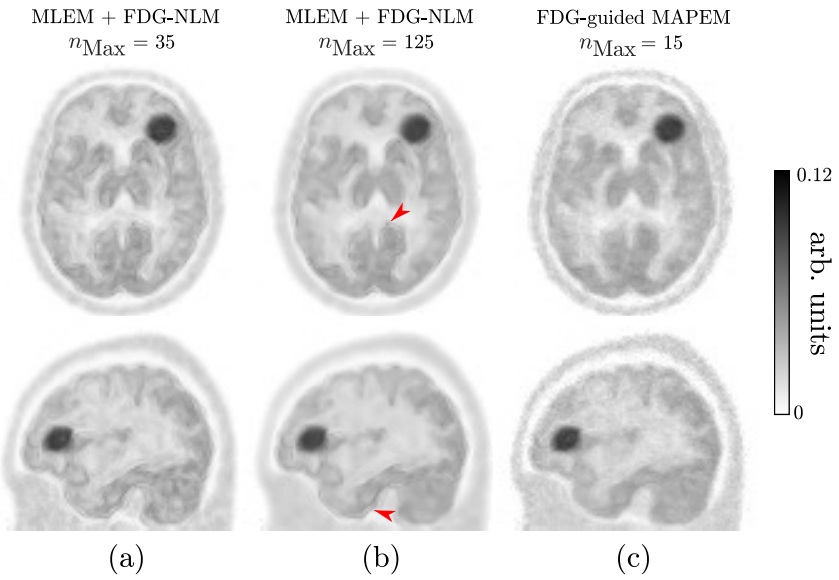


Figure 5.9: Example reconstructed images from the investigation into the effect of the hyperparameter n_{Max} on image quality for FDG-NLM filtered MLEM and FDG-guided MAPEM. (a) FDG-NLM filtered MLEM with $n_{\text{Max}} = 35$, (b) FDG-NLM filtered MLEM with $n_{\text{Max}} = 125$, and (c) FDG-guided MAPEM with $n_{\text{Max}} = 15$. Note that with a lower value of n_{Max} (i.e. less computational burden), the FDG-guided MAPEM produces images of equal or superior quality compared to the FDG-NLM filtered MLEM images. Arrows indicate examples of the FDG-NLM artefacts which remain even with $n_{\text{Max}} = 125$.

value of either σ_f or σ_s. The performances of both FDG-NLM filtered MLEM and FDG-guided MAPEM at high σ_f and σ_s are similar to their performances as shown in Figure 5.6.

When the [^{18}F]-FDG dataset was reconstructed with 50 or 200 iterations of MLEM, instead of the 100 used above, the resultant prior images were visually similar (Figure 5.11). The 50-iteration [^{18}F]-FDG prior (FDG_{50}) displays lower levels of noise than the 100-iteration [^{18}F]-FDG prior (FDG_{100}) and some blurring from being under-converged, and the 200-iteration prior (FDG_{200}) exhibits more noise, although this is mitigated by the subsequent 4 mm Gaussian smooth. The resulting FDG-guided [^{11}C]-methionine reconstructions all appear visually similar, and the difference between them and the FDG_{100}-guided image is observed to be small, with a maximum difference in observed tumour mean of 0.44%. This difference was smaller for the FDG_{200}-guided [^{11}C]-methionine reconstruction, indicating that the difference

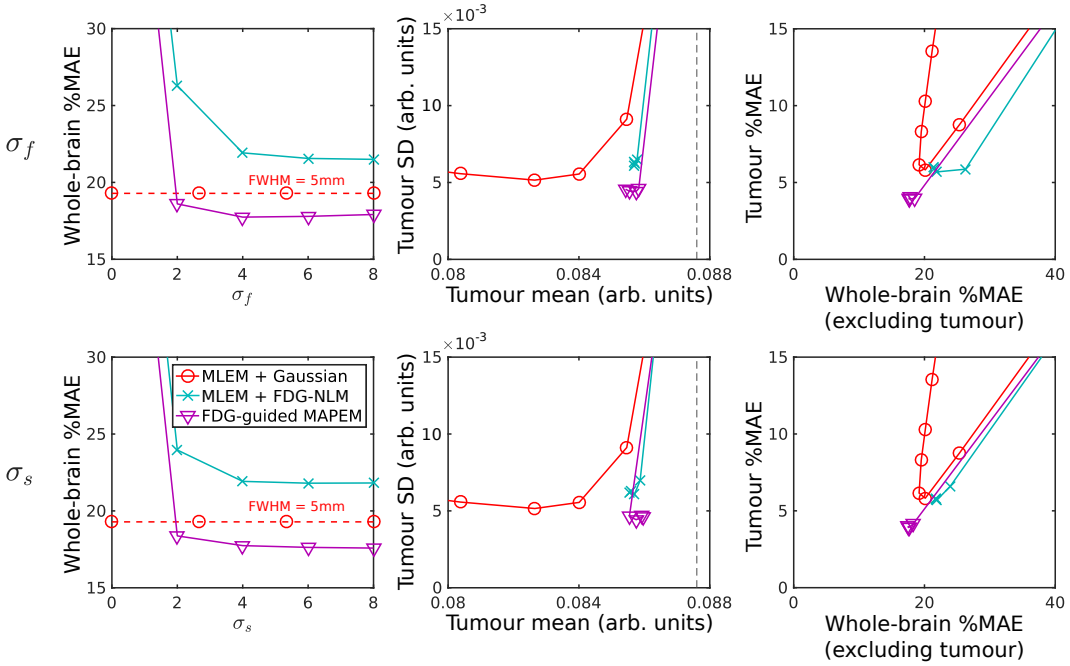


Figure 5.10: Effect of the hyperparameters σ_f (left) and σ_s (right) on image metrics for FDG-guided MAPEM and FDG-NLM filtered MLEM, with all other hyperparameters set to their baseline values as listed in Table 5.1. Also shown is Gaussian filtered MLEM. At $\sigma_f \geq 2$ and $\sigma_s \geq 2\,\text{mm}$, FDG-NLM filtered MLEM and FDG-guided MAPEM perform similarly to the results seen in Figure 5.6. © 2018 IEEE

in calculated weights is smaller when moving from 100 to 200 iterations for the prior compared to moving from 50 to 100 iterations.

Finally, when the tumour was changed for a [^{11}C]-methionine-specific tumour in the mismatch test, performing FDG-guided [^{11}C]-methionine image reconstruction was observed to attenuate the observed tumour intensity to a greater extent than that which was observed when using an MLEM reconstruction with a 5 mm Gaussian smooth (Figure 5.12). This manifested as a 4.3% reduction in tumour to white matter CRC (from 1.84 using smoothed MLEM to 1.76 using FDG-guided MAPEM).

5.3 Real Data Preliminary Study

Following the simulation study, the proposed FDG-guided image reconstruction method was applied to a real [^{18}F]-FDG/[^{11}C]-methionine scan pair from a single suspected glioma patient who was scanned in a Siemens Biograph mMR

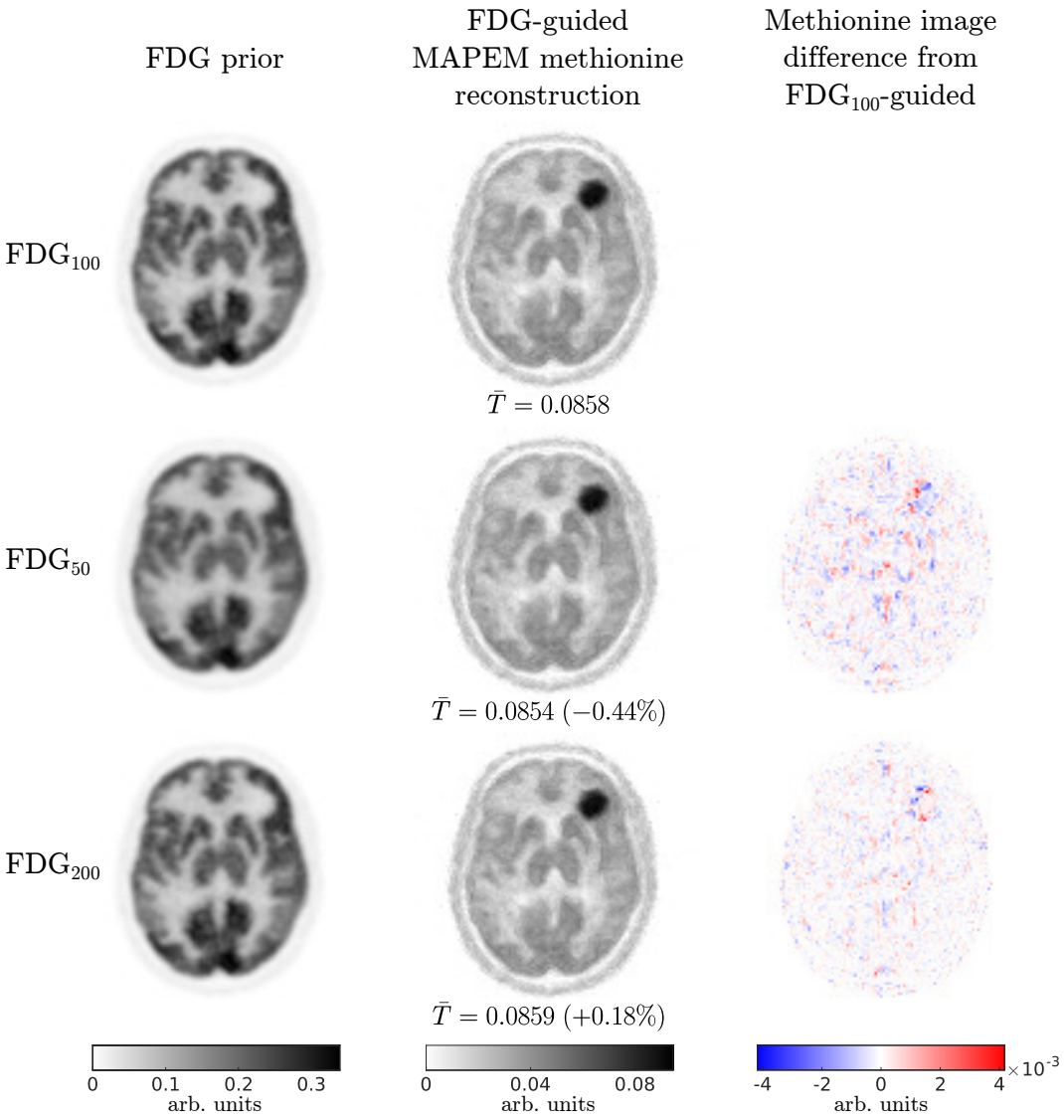


Figure 5.11: The effect of changing the reconstruction of the [^{18}F]-FDG image prior on the FDG-guided MAPEM [^{11}C]-methionine reconstructions in simulated data. Left column: 4 mm smoothed [^{18}F]-FDG priors with 100, 50, and 200 iterations of MLEM (FDG_{100}, FDG_{50}, and FDG_{200} respectively); middle column: the corresponding FDG-guided [^{11}C]-methionine reconstructions; and right column: the difference between each [^{11}C]-methionine reconstruction and the FDG_{100}-guided [^{11}C]-methionine reconstruction. $\bar{T}$ denotes the measured tumour mean $\text{Mean}_T(\boldsymbol{\theta})$, and values in brackets denote the percentage deviation from the FDG_{100}-guided tumour mean measurement.

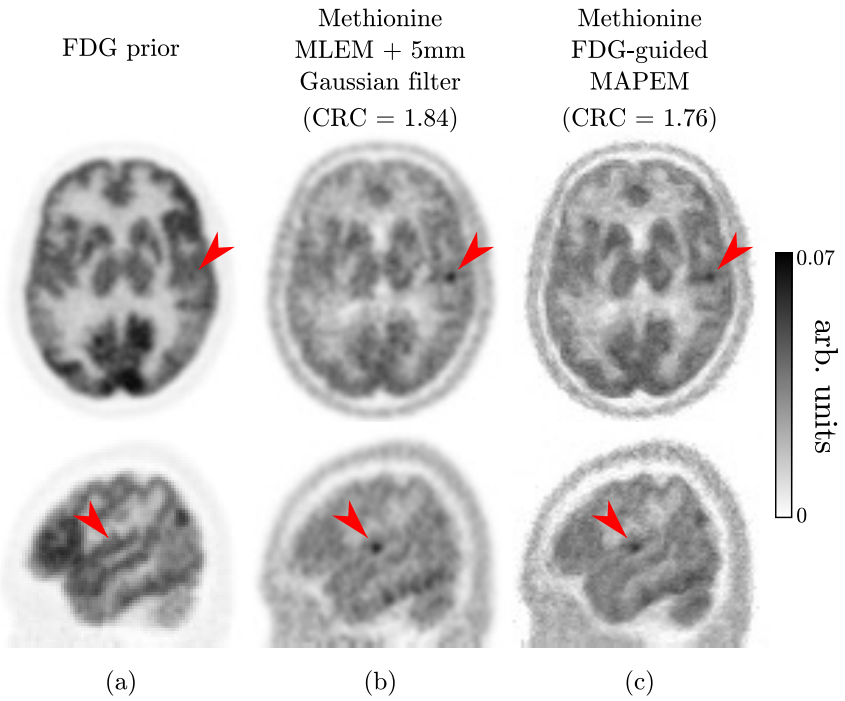


Figure 5.12: Reconstructed images for the mismatch test. (a) the [^{18}F]-FDG prior, reconstructed with 100 iterations of MLEM followed by a 4 mm Gaussian smooth; (b) the [^{11}C]-methionine data reconstructed with 200 iterations of MLEM followed by a 5 mm Gaussian smooth; and (c) FDG-guided [^{11}C]-methionine reconstruction using the baseline hyperparameter values as listed in Table 5.1. Arrows indicate the location of the [^{11}C]-methionine-specific tumour, which was attenuated more by the FDG-guided reconstruction than the Gaussian smoothed MLEM reconstruction. This was reflected by a 4.3% reduction in tumour to white matter CRC. © 2018 IEEE

PET-MR scanner (Siemens Healthcare, Erlangen, Germany). The patient was scanned using both radiotracers on the same day. The [^{11}C]-methionine scan was performed first, with an injected activity of 337 MBq and a scan duration of 11.4 min, beginning at 40 min post-injection. Fifty-six minutes after the [^{11}C]-methionine injection, 233 MBq of [^{18}F]-FDG was injected into the patient, who was then scanned for 11.7 min at 54 min post-injection. The [^{18}F]-FDG dataset contained ∼381 M counts, and the [^{11}C]-methionine dataset contained ∼73.7 M. The vendor-provided e7 Tools software (Siemens Healthcare, TN, USA) was used to obtain normalisation factors and scatter and randoms estimates for these datasets. Attenuation factors were calculated using the vendor-provided ultra-short echo time MR-based attenuation maps.

5.3.1 Experimental Methods

The FDG-guided [^{11}C]-methionine image reconstruction pipeline for the real data is shown in Figure 5.13. The [^{18}F]-FDG dataset was first reconstructed into the standard $344 \times 344 \times 127$ image grid with 100 iterations of MLEM followed by a 4 mm Gaussian smooth, using the same in-house geometric projectors used for the 3D simulation study. Following this, the [^{18}F]-FDG image was transformed to the [^{11}C]-methionine image space via a rigid-body alignment transformation estimated from the simultaneously acquired MR-based attenuation maps. The transformed [^{18}F]-FDG image was then used to calculate the set of weights w_{jk} as described above. Here the baseline hyperparameter values from the hyperparameter experiment results were used, as listed in Table 5.1, due to the overall good reconstruction performance observed using these values. These weights were then input into the reconstruction of the [^{11}C]-methionine dataset (i.e. 200 iterations of Equation 2.36) to produce the output FDG-guided [^{11}C]-methionine images. For comparison purposes, two additional sets of guidance weights were calculated from the co-acquired T1-weighted MR image and a previously acquired aligned CT image using the same hyperparameter values, and the [^{11}C]-methionine dataset reconstructed using these weights. Note that resolution modelling was not included in these reconstructions.

Finally, a validation experiment was carried out on the real data in order to investigate the extent to which guiding a [^{11}C]-methionine image reconstruction with an [^{18}F]-FDG image affects the structure of the output [^{11}C]-methionine image. Ideally this would be achieved by using a high-counts [^{11}C]-methionine dataset to produce a reference image against which lower counts reconstructions using FDG-guided MAPEM or MLEM could be compared. However, due to the counts-limited nature of the available [^{11}C]-methionine dataset, this experimental approach was not possible. Instead, a similar methodology was employed whereby a low-noise [^{11}C]-methionine image was produced by reconstructing the data with 200 iterations of MLEM into a coarse image grid ($115 \times 115 \times 42$) at three times the standard mMR voxel size listed above.

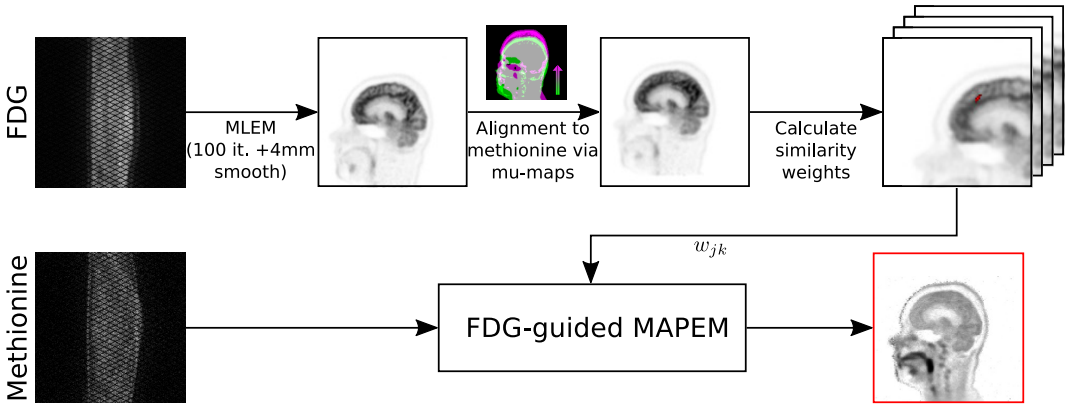


Figure 5.13: Data processing pipeline used for the FDG-guided image reconstruction of the real [^{11}C]-methionine dataset. First, the [^{18}F]-FDG image was reconstructed with 100 iterations of MLEM and then smoothed with a Gaussian kernel of FWHM 4 mm. This reconstructed image was then transformed to the [^{11}C]-methionine image space using rigid-body transformations estimated from the attenuation maps, before the similarity weights w_{jk} were calculated. The final step involved inputting these weights into the MAPEM reconstruction algorithm alongside the measured [^{11}C]-methionine data to produce FDG-guided MAPEM reconstructions. © 2018 IEEE

Counts-reduced [^{11}C]-methionine datasets were then generated by randomly removing counts from the original [^{11}C]-methionine sinogram according to a Bernoulli process, with the probability of removal equal to the expected proportion of counts remaining. In this way, [^{11}C]-methionine datasets with 10% and 20% the original data were generated. The original [^{18}F]-FDG dataset was reconstructed into the same coarse image grid to calculate the weights for FDG-guided MAPEM. The counts-reduced [^{11}C]-methionine datasets were then reconstructed with 200 iterations of FDG-guided MAPEM and Gaussian filtered MLEM. Reconstruction error was quantified by measuring the %MAE between reconstructed counts-reduced images and the original 100% MLEM reconstruction in brain voxels within a single representative image slice.

5.3.2 Results

Figure 5.14 shows the reconstructed images for the real data case, including the aligned [^{18}F]-FDG prior image, the T1-weighted MR image acquired simultaneously with the [^{11}C]-methionine data, and the aligned CT image acquired prior to the PET-MR scan. Note that the MR and CT images have been resampled

to the PET image grid, and that the CT has been contrast-enhanced using a 3D contrast limited adaptive histogram equalisation. It can be seen that the [^{11}C]-methionine image reconstructed using FDG-guided MAPEM is visually sharper than the 5 mm Gaussian filtered MLEM reconstruction, in agreement with the 3D simulation study results. Furthermore, the FDG-guided MAPEM reconstruction is sharper than both the CT- and MR-guided MAPEM reconstructions. This is particularly visible in the thalamus (Figure 5.14(a), filled red arrows and Figure 5.14(c)), where the FDG-guided MAPEM image exhibits a higher regional uptake than Gaussian filtered MLEM and both CT- and MR-guided reconstructions. Importantly, when considering a region where there is a discrepancy between the [^{18}F]-FDG and the [^{11}C]-methionine images in the anterior temporal lobe (Figure 5.14(b), unfilled red arrows), the [^{11}C]-methionine-specific uptake pattern was preserved in the FDG-guided MAPEM images.

When β is varied from 2000 to 18000 (Figure 5.15), the effects of the penalty strength can be more clearly observed. There does not appear to be any change in the visual appearance of the radiotracer distribution as β increases, only an improvement in the noise characteristics of the images.

For the validation study using the coarse reconstructed image grid, the guidance hyperparameter values were set based on the baseline values used above. Since the voxel size increased by a factor of three, features were reduced from $3 \times 3 \times 3$ patches to single voxel values and the neighbourhoods were reduced from $5 \times 5 \times 5$ to $3 \times 3 \times 3$. σ_f was left equal to 4 since features are normalised, and σ_s was set to 8 mm to allow suitably large similarities between neighbouring voxels. n_{Max} was reduced from the previous baseline value of 15 to 3 in order to keep the same proportion of voxels in the weights ($15/125 \rightarrow 3/27$), and β was varied between 2.5×10^5 and 3×10^7 since redefining the voxel grid scales the value of the log-likelihood term.

Figure 5.16(a) shows the brain %MAE in the analysed slice as a function of β, demonstrating that the optimum β value depends on the counts level of the dataset. Similarly, for Gaussian smoothed MLEM (Figure 5.16(b)), the optimum smoothing level depends on the counts level. It can be seen that

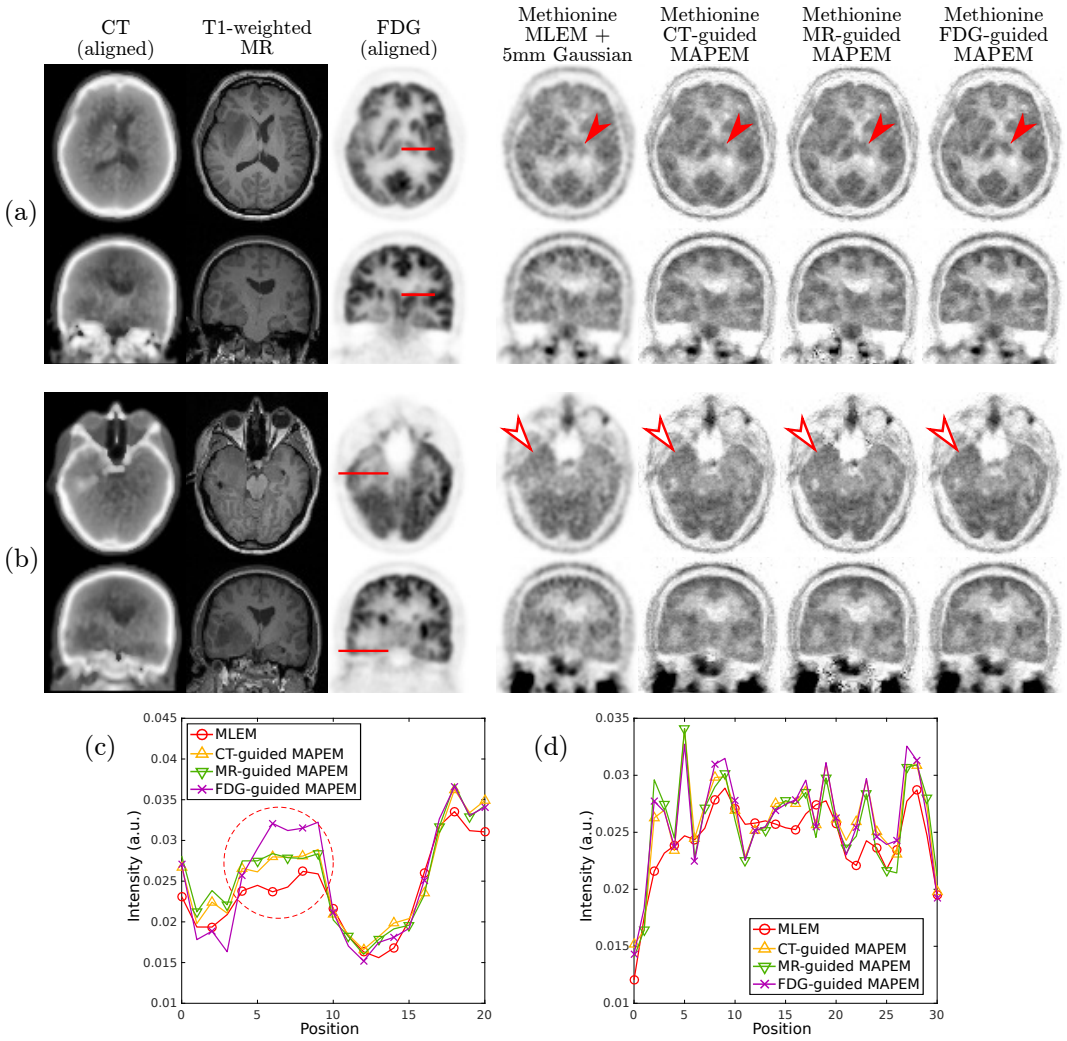


Figure 5.14: Reconstructed images for the real data preliminary study. From left to right for both (a) and (b): the contrast-enhanced aligned CT image; the T1-weighted MR image; the [^{18}F]-FDG data reconstructed with 100 iterations of MLEM with a 4 mm Gaussian smooth; the [^{11}C]-methionine data reconstructed with MLEM followed by a 5 mm Gaussian smooth; the [^{11}C]-methionine data reconstructed with CT-guided MAPEM; [^{11}C]-methionine data reconstructed with MR-guided MAPEM; and the [^{11}C]-methionine data reconstructed with FDG-guided MAPEM. (c) and (d) show the [^{11}C]-methionine image profiles for (a) and (b) respectively, with the locations of the profiles indicated on the [^{18}F]-FDG images. All [^{11}C]-methionine reconstructions were run to 200 iterations, and all guided reconstructions used the baseline hyperparameter values as given in Table 5.1. (a) Increased uptake was observed in the thalamus when using the FDG-guided MAPEM compared to the Gaussian filtered MLEM and CT- and MR-guided reconstructions (filled arrows and circled in the profile). (b) A discrepancy between [^{18}F]-FDG and [^{11}C]-methionine uptakes is apparent (unfilled arrows), however the FDG-guided MAPEM allows good reconstruction of this region without any visible cross-talk.

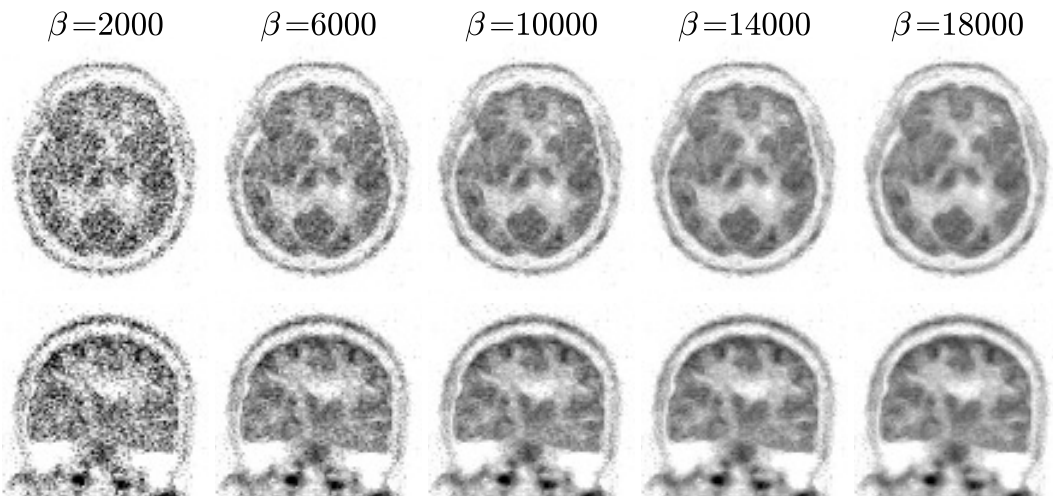


Figure 5.15: Real data FDG-guided MAPEM reconstructions as a function of β, using the same weights hyperparameters as were used in Figure 5.14. Noise levels decrease as β increases, without visually affecting the pattern of distribution of the radiotracer. © 2018 IEEE

using FDG-guided MAPEM with 10% counts allows for %MAE values similar to those obtained with MLEM using 20% counts. When considering the images at the best %MAE values for each method/counts level combination (Figure 5.16(c)), FDG-guided MAPEM is observed to produce images that are visually less noisy than the counts-matched MLEM reconstruction. Furthermore, the structure in the FDG-guided MAPEM reconstruction is seen to agree with that present in the 100% counts image. Finally, the error maps obtained with the FDG-guided MAPEM (Figure 5.16(d)) do not show any visible structure, supporting the observation that the use of the FDG-guidance did not introduce any visual artefacts into the reconstructed images.

5.4 Discussion

This Chapter aimed to evaluate the novel proposal of using guidance information from one PET image in the reconstruction of another PET dataset produced using a different radiotracer. To evaluate the possible benefits of such a methodology, a 3D [^{18}F]-FDG/[^{11}C]-methionine brain tumour simulation study was performed to explore the hyperparameter space of weighted quadratically-penalised FDG-guided reconstruction. The resulting reconstructed images

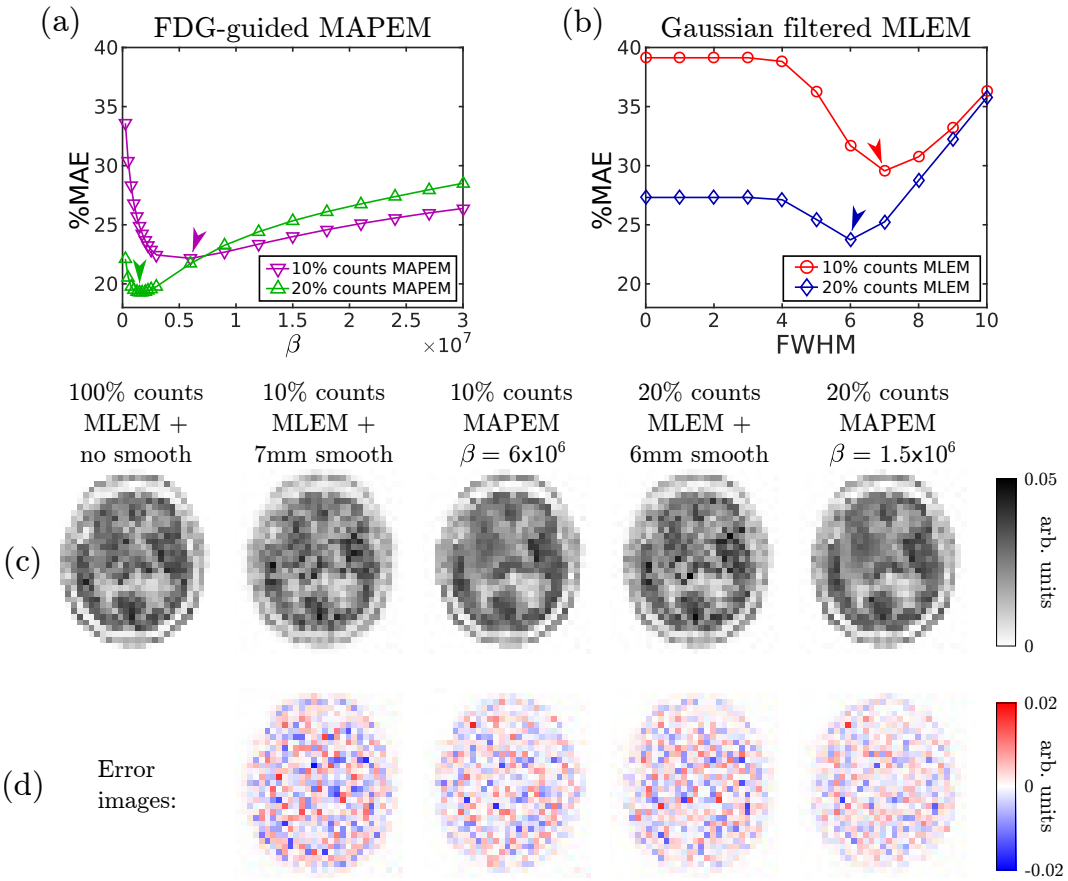


Figure 5.16: Results of the validation experiment using real data reconstructed at a larger voxel size that was 3× the standard mMR voxel size. (a) Reconstruction %MAE in the analysed slice relative to the 100% MLEM reference image as a function of β for FDG-guided MAPEM at both 10% and 20% counts levels, (b) the corresponding %MAE values for MLEM with a varying Gaussian smooth, (c) reconstructed images at the best hyperparameter value for FDG-guided MAPEM and MLEM for the two counts levels, with FDG-guided MAPEM exhibiting lower image noise than the corresponding MLEM reconstructions, and (d) the error maps for the the images in (c), showing that using FDG-guided MAPEM does not appear to introduce any structural artefacts into the error image.

were compared to MLEM smoothed with either a post-reconstruction Gaussian smooth or an [^{18}F]-FDG-derived NLM filter. Following this, a real [^{18}F]-FDG/[^{11}C]-methionine dataset pair were reconstructed with the proposed method and compared to post-reconstruction filtered MLEM and CT- and MR-guided MAPEM images.

In terms of the penalty strength β, it was observed that FDG-guided MAPEM outperformed post-reconstruction filtering methods in terms of reconstruction error in both the tumour and in the rest of the brain, indicating that FDG-guidance in a quadratic penalty achieves the noise reduction and edge-preservation that is desired. With increasing values of β, whole-brain error reduced (Figure 5.6(a)), and was still reducing at the maximum β value tested (25000), implying that there is potentially a higher β value than provides a lower whole-brain %MAE. Nevertheless, at $\beta = 25000$, the tumour mean value had already become considerably biased (Figure 5.6(b)), indicating that a higher β value, while more optimal across the whole-brain, would be sub-optimal in the tumour compared to the other values tested in this work.

An unexpected observation from the results of varying β was that the FDG-NLM filtered MLEM reconstruction performed worse than applying a 5 mm Gaussian filter in terms of whole-brain error, despite the use of information from the [^{18}F]-FDG image in creating the NLM filter. The reason for this lies in the support of the filters; the [^{18}F]-FDG-derived NLM filter had only $n_{\text{Max}} = 15$ non-zero elements, whereas the Gaussian filter had many more non-zero elements. Therefore, the noise reduction achieved by the Gaussian filters was superior to that obtained by using the FDG-NLM filter. This is supported by Figure 5.8, where using the FDG-NLM filter with a higher value of n_{Max} produced reconstruction errors lower than those provided by the 5 mm Gaussian filter.

Conversely, since the FDG-guided MAPEM provides regularisation during the reconstruction process, smoothing is effectively applied at each iteration (according to Equation 2.37) and so can propagate over the course of many iterations. This ensures that sufficient noise reduction can be achieved with a lower value of n_{Max}, without compromising edges or introducing the artefacts

observed in the FDG-NLM reconstructions in Figures 5.7 and 5.9. It is possible that alternative methods of using the same [^{18}F]-FDG guidance information, such as inter-iteration NLM filtering, or the use of [^{18}F]-FDG-derived spatial basis functions (c.f. [181,182,185]), could alleviate the problems observed when using NLM filters, although such a comparison was beyond the scope of this work.

While β and n_{Max} were seen to have a strong impact on reconstructed image quality, the hyperparameters σ_f and σ_s were observed to have a relatively small impact (Figure 5.10). At sufficiently high values of each, the whole-brain %MAE, tumour mean and SD, and tumour %MAE were all stable, with results reflecting those seen in Figure 5.6. This may be due to the baseline values chosen for the other hyperparameters, particularly n_{Max}. Since the baseline value of n_{Max} was relatively low (15), the effect of the two kernel widths was limited; that is to say that once the maximum 15 values had been selected for each voxel, they were sufficiently similar to that voxel that widening the weights kernels would make no difference. Future work could account for this by repeating the analysis of the effects of σ_f and σ_s at various values of n_{Max}.

The reconstruction hyperparameters of the [^{18}F]-FDG prior were also observed to have a relatively weak impact on the quality of the FDG-guided reconstructions (Figure 5.11). The resultant [^{11}C]-methionine images appeared visually similar when using FDG priors reconstructed with 50, 100, or 200 iterations of MLEM followed by a 4 mm Gaussian smooth. This suggests that as long as the guiding image is of sufficient quality (i.e. has a sufficient number of measured counts), the calculated weights will be similar, since the Gaussian smoothed MLEM image estimate will converge to a low-noise solution. Additional work could be carried out to assess more fully this impact; in particular it would be important to know the minimum number of FDG counts that allow for reliable [^{11}C]-methionine image improvement in a given application before using the resulting FDG-guided [^{11}C]-methionine images in place of standard MLEM images.

Future work will also require hyperparameter optimisation based on a larger number of patient datasets, evaluated in an application-specific manner with

figures of merit and image quality metrics that relate more directly to the clinical need of the multi-tracer protocol in question. While %MAE and other error-based metrics are useful for the preliminary evaluation of reconstruction methods in simulation studies, they can tend to prefer oversmoothed images (as seen in Figures 5.8 and 5.9), and so can suggest a sub-optimal set of reconstruction hyperparameters for a given clinical problem.

Nonetheless, when demonstrated on real data and compared to CT- and MR-guided reconstructions, the FDG-guidance performed well, visually improving the sharpness and contrast of the thalamus (Figure 5.14). This is an important result as it suggests that there is some additional/higher quality information present in the [^{18}F]-FDG image that is not present in either the CT or MR images. This is supported by visual inspection of the [^{18}F]-FDG, CT, and MR images, where the inner brain structures possess a visibly higher contrast in the [^{18}F]-FDG image. However, it is important to consider that CT and particularly MR images generally possess a superior resolution to PET images. By not including resolution modelling in the reconstructions in this Chapter, the performances of the CT- and MR-guided reconstructions may be degraded compared to their optima.

In general, when performing guided PET image reconstructions, care must be taken to avoid the so-called 'cross-talk' problem: where smoothing penalties are incorrectly translated from the guiding image into the image to be reconstructed, giving rise to the appearance of features that are not present in the true radiotracer distribution such as false edges. In the hyperparameter characterisation study performed in this Chapter, the same anatomical boundary locations were used for both the [^{18}F]-FDG and [^{11}C]-methionine ground truth images (Figure 5.3), potentially unfairly benefiting the FDG-guided MAPEM since there was a true close correspondence between the two images. In fact, when a mismatch was introduced into the phantom by inserting a small [^{11}C]-methionine-specific tumour, this feature was blurred more in the FDG-guided image reconstruction than in the standard Gaussian smoothed MLEM reconstruction (Figure 5.12). This result suggests that although the proposed method can perform well in some situations, it is inadvisable to

use exclusive FDG-guidance in a weighted quadratic penalty when such mismatches are possible. Therefore, before widespread use on clinical data, the application-specific hyperparameter characterisation mentioned above would be vital to ensure that the risk of attenuating [^{11}C]-methionine-specific features is limited. Alternatively, more advanced guidance methods could be utilised that would allow the [^{11}C]-methionine data to override the [^{18}F]-FDG information when appropriate (e.g. [170]). On the other hand, the fact that the images reconstructed from the real data display improvements with no visible cross-talk in regions where the radiotracer distributions differ (Figure 5.14) indicates that FDG-guided MAPEM can be robust to the differing distributions of uptake for two radiotracers that correspond to distinct biological processes. This robustness was validated further by performing a counts reduction experiment whereby [^{11}C]-methionine reconstructions at 10% and 20% counts levels were compared to the MLEM reconstruction of the 100% counts dataset. However, due to the low counts levels in the original dataset, validation had to be performed at a larger voxel size to provide sufficient counts per voxel in the 100% counts reference image. Nevertheless, the results showed reduced reconstruction error relative to this reference image when using FDG-guided MAPEM compared to MLEM. The reconstructed images and the error maps did not show any visible structure that would be indicative of cross-talk between the datasets, reinforcing the indications that the method is robust to cross-talk issues at least down to this voxel size. It would useful to repeat this experiment in future with a high-counts [^{11}C]-methionine dataset that allows for the same analysis to be repeated at the standard clinically used voxel size to verify that there are no smaller radiotracer features that are degraded by reconstructing [^{11}C]-methionine images with FDG-guidance.

5.5 Summary

In this Chapter we proposed the use of higher quality PET images to guide the reconstruction of lower quality PET images acquired using other radiotracers. By using the high quality PET image to define which sets of voxels in the lower

quality PET image are expected to have similar intensities, edge-preserving, contrast-invariant, regularised image reconstruction can be performed in a manner similar to that used in anatomically guided PET reconstruction. To evaluate the benefits of such a methodology, this Chapter focused on the specific case of glioma imaging, where using images from both [^{18}F]-FDG and [^{11}C]-methionine can be used to improve diagnosis. High quality [^{18}F]-FDG images were used to derive inter-voxel similarity weights within neighbourhoods to guide a quadratically-penalised MAPEM reconstruction of the [^{11}C]-methionine datasets. 3D simulation studies demonstrated lower reconstruction errors in [^{11}C]-methionine images reconstructed using FDG-guided MAPEM, as well as superior quantification in a tumour that was hypo-intense in terms of [^{18}F]-FDG uptake and hyper-intense in terms of [^{11}C]-methionine uptake. When applied to real patient data, similar improvements were seen across the image, and furthermore, the [^{11}C]-methionine-specific radiotracer distribution was maintained when using [^{18}F]-FDG guidance.

Future work will involve further validation of the proposed methodology on greater numbers of real dataset pairs and extension of the methodology to other radiotracer pairings. In addition, other penalties from the anatomically guided PET image reconstruction literature could be tested in the multi-tracer PET context, in order to attempt to find optimum methods for sharing information between datasets of complementary radiotracers.

Chapter 6:

Discussion

This thesis has presented novel methods for sharing information between multiple PET datasets in order to improve the quality of reconstructed images. In this Chapter, a general discussion about the presented methods is provided, including limitations, potential applications and areas of future work.

6.1 Difference-Image Priors for Simultaneous Reconstruction of Longitudinal Datasets

In Chapters 3 and 4, image reconstruction priors that acted in the difference image domain were defined in order to improve longitudinal PET images for tumour monitoring. Each of these priors penalised different aspects of the difference image, according to mathematical formulations of the prior expectations on the appearance of such difference images. The assumptions and properties of each of these priors are briefly reviewed here, before describing what would be desirable from an ideal difference-image prior. Table 6.1 summarises the observed effects of the proposed priors and their main characteristics, while Figure 6.1 shows the effect of the penalty strength β on difference images for a simplified 2D case reconstructed using each prior.

Firstly, the DS priors were designed to encourage reconstructed longitudinal images that resulted in sparse difference images. As long as the majority of the background voxels in the corrected difference image (Equation 4.2) are zero or below some noise threshold, this method achieves meaningful sparsification,

	DS	DE	DTV
Characteristic encouraged	Sparsity	Low entropy	Sparse spatial gradients
Appearance of difference images*	Predominantly zeroes	Piecewise constant	Piecewise constant
Contrast-invariant?†	No	Partially	Yes
Noise reduction in tumours	Low	Medium	High
Benefits	Preserves MLEM quality for tumours	Low bias in tumours	Low %RMSE in tumours
Limitations	Limited image improvement in tumours	Sensitive to hyperparameters	Higher tumour bias

* see Figure 6.1 for examples.
† i.e. invariance to spatially dependent scaling factors between scans, e.g. grey matter and white matter changing by differing factors.

Table 6.1: Observed characteristics of the proposed difference-image priors. Note that some characteristics can be an advantage or a drawback depending on the context. For example, the preservation of MLEM image quality in tumours produced by the DS prior may be an advantage if strict adherence to the data is prioritised, but may be a drawback in terms of the lack of scope for improvement of tumour reconstructions.

as seen previously in Chapters 3 and 4. As the penalty strength increases, the number of non-zero elements in the difference image decreases (Figure 6.1). In cases where the corrected difference is not *a priori* sparse though, such as when grey and white matter have separate scaling factors between scans, the DS prior would be expected to perform poorly due to the existence of additional valid non-zero differences throughout the image. In the experiments performed in this thesis, the DS priors achieved higher levels of noise reduction in the background, while a lower level of noise reduction was observed in tumours, resulting in images in which the tumours appeared visually similar to the standard MLEM reconstructions. This property could be considered a benefit of the DS prior: such tumours represent minimal imposition of prior expectations where there is no additional data to support noise reduction. However, this means that the improvement in image quality is limited in the

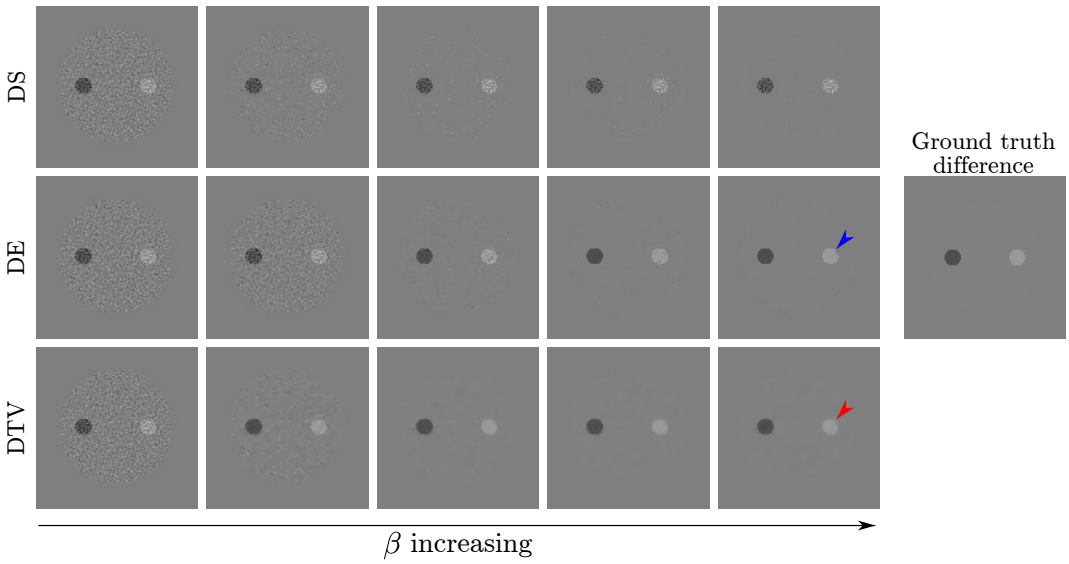


Figure 6.1: Visualisation of the effect of the DS, DE, and DTV priors using a simplified 2D example. Top row: As the penalty strength, β, increases, difference images contain a greater number of zero values. Middle row: The DE prior provides piecewise constant difference images with sharp edges around regions of change (blue arrow). Bottom row: The DTV prior also produces piecewise constant difference images, but with a spatial blurring (red arrow).

tumours, i.e. the areas of interest.

The DE prior seeks image estimates that corresponded to low entropy difference images. Since low entropy can be visualised as sharp peaks in the image histogram, corresponding to distinct, separated sets of voxel intensities, the DE prior can be interpreted as trying to represent the difference image in terms of the fewest possible unique differences. Although this property can, in general, manifest in a spatially disjoint manner, using the DE prior with typical change and noise levels tends to create uniform regions to achieve this, as observed in Figure 6.1. In particular, in can be seen that the DE prior provides sharp edges, even at high penalty values (blue arrow). This method is partially invariant to contrast changes between images, as long as these changes do not greatly increase image heterogeneity. For the example mentioned above, where grey and white matter uptake values scale separately but consistently between scans, the DE prior could perform better than the DS prior by separating the grey and white matter peaks in the difference image histogram. However, in cases of low-amplitude, heterogeneous contrast changes, (e.g. a small heterogeneous increase in grey matter activity) the DE

prior may over-regularise reconstructions in order to merge voxels into a single difference image peak. In terms of noise reduction within the tumours, the DE method was observed to have an intermediate effect, caused by its tendency to group voxels into peaks. The bias of tumour mean values was low with the DE method in 2D simulations, although this behaviour was not observed in the real data experiment.

Finally, the DTV prior aimed to produce image estimates that resulted in difference images with sparse spatial gradients. This achieves a similar outcome to the DE prior in terms of encouraging uniform regions in the difference image (Figure 6.1). However, the visual appearance of reconstructed difference images is different due to the form of the DTV prior; because the spatial gradients of the difference image are used, there is an explicit spatial regularisation that is absent from the DE prior. This manifests as a blurring in regions of change (Figure 6.1, red arrow) and is likely to be the cause of the increased biases that were observed for the DTV prior in tumours with small radii (e.g. Figures 4.4(b) and 4.12). However, within tumour regions in the reconstructed images, the DTV prior allowed greater noise reduction compared to the DE prior. Indeed, the DTV prior produced the lowest tumour %RMSE values of the three proposed difference-image penalties (Figures 4.5 and 4.15).

6.1.1 Improving the Difference-Image Priors

Considering the previous discussion about the characteristics of the priors proposed in Chapters 3 and 4, a description of the characteristics of an ideal difference-image prior may be attempted, along with suggestions about how this ideal prior may be achieved.

The ideal prior depends on the characteristics expected from the images in question. To allow the definition of such expectations, this Section considers difference images in which the only changes present come from the tumours, and these tumours have homogeneous uptake. Using these assumptions allows the definition of the following characteristics of an ideal difference-image prior, in order of preference:

1. Background noise reduction equivalent to counts-sharing between all scans,
2. Preservation of tumour mean intensities, allowing longitudinal comparisons of mean SUV, and
3. Noise reduction within the tumour, lowering variability of tumour quantification.

Requirement 1 is met by all of the proposed difference-image priors, as shown in the simulation study in Chapter 4. Requirements 2 and 3 are achieved by both the DE and the DTV priors to differing degrees. For example, the DE prior best preserved tumour mean values, whereas the DTV prior provided the highest levels of noise reduction within tumours.

Of the three methods, the DE prior was the optimum one in simulation studies, as its only drawback was a relative lack of noise reduction compared to the DTV prior. Furthermore, the amount of noise reduction achievable using the DE prior is likely dependent on the Parzen window width used. While narrower Parzen windows would allow a greater degree of noise reduction by being able to better discriminate similar difference values, the use of the OSL reconstruction algorithm limited the Parzen window widths that could be used in this work. Alternative image reconstruction algorithms, such as the preconditioned conjugate gradient algorithm used in [178], could be used to allow narrower Parzen windows for enhanced tumour noise reduction.

However, although the DE prior was optimal in 2D simulation studies, real data results indicated that the DE prior performed worse than the competing DS and DTV methods in terms of tumour bias (Figure 4.16). This is likely the result of the presence of other changes in the difference image that affected the DE prior due to its global nature. Specifically, the existence of changes with similar magnitudes elsewhere in the difference image caused an increased bias, since the DE prior attempted to merge these spatially distinct regions into a single peak in the difference image histogram in order to reduce its entropy.

This limitation can be alleviated by including spatial information in the entropy-based prior. There are several approaches for including such information. One approach is to generate a set of multi-scale images that capture the

desired spatial information (e.g. with a Gaussian pyramid), and define the entropy in terms of this set of images [178, 179]. Another approach is to use two penalty terms, one entropy-based prior and one spatial prior in order to achieve both types of regularisation simultaneously [176]. This latter method could be easily incorporated into the difference-image priors by using both the DE and DTV priors simultaneously.

Another method of including spatial information into entropy-based priors is to calculate entropy in smaller regions throughout the image rather than across the whole image, and then aggregate these local entropy measures. For example, the localised entropy (LE) of an image $\boldsymbol{\theta}$ could be defined as

$$\mathrm{LE}(\boldsymbol{\theta}) = \sum_{j=1}^{J} \hat{H}_j(X, \boldsymbol{\theta}), \tag{6.1}$$

where

$$\hat{H}_j(X, \boldsymbol{\theta}) = \hat{H}(X, \{\theta_k : k \in \mathcal{N}_j\}) \tag{6.2}$$

is the entropy of the voxel intensities within the image neighbourhood $\mathcal{N}_j$. Because entropy is non-linear, the sum of the entropy in all neighbourhoods is distinct from the entropy of the entire image.

Using the LE of difference images would aim to produce difference images that possess low entropy in neighbourhoods, rather than in the entire image. In this way, spatially distinct regions would be prevented from affecting each other.

Figure 6.2 shows a simplified 2D example, demonstrating the potential improvements that could be obtained by using the local DE (LDE) prior compared to the global DE prior. In this case there are three regions of change, two of which change by a similar, but not identical, amount between the PET1 and PET2 images (labelled A and B). When images are reconstructed with the DE prior, the two regions A and B are merged into a single peak in the difference image (labelled C). This corresponds to an induced bias in both the A and B regions of change. However, the LDE prior is able to maintain the separation between the peaks, while achieving comparable levels of noise

reduction in both the reconstructed images and in the difference image.

These preliminary results suggest that the LDE prior may be able to improve on the DE prior in cases like that observed in Figure 4.17, but a comprehensive simulation study such as that performed in Chapter 4 is required to fully characterise its performance. Furthermore, the LDE prior is considerably more computationally expensive than the DE prior as it requires calculation of the DE prior on all neighbourhoods, necessitating the use of more optimal algorithms to reduce reconstruction times. One potential method for increasing the LDE-PML reconstruction speed is to use Fourier methods to more efficiently estimate the required difference image PDFs [178].

It is worth noting that although the LDE prior may be a promising area of future work, it is still limited by the assumption of tumour homogeneity, and the focus on mean uptake as a marker of tumour activity. There are many cases where longitudinal tumour monitoring difference images may be highly heterogeneous, for example when the tumours themselves are heterogeneous [241], or when there has been a heterogeneous response to therapy [88]. Similarly, other measures of tumour activity, such as maximum tumour SUV, may be of interest [70, 79, 80]. In these cases, the DTV prior may be preferable since it has been observed to better preserve the pattern of heterogeneous regions (c.f. Figure 4.18). The evaluation of difference-image priors for such heterogeneous cases remains for future work.

6.1.2 Relation to the Existing Multi-Dataset Image Reconstruction Literature

To the best of our knowledge, the concept of simultaneously reconstructing longitudinal oncology treatment monitoring PET datasets has not been proposed before in the research literature. In this thesis, this was achieved by using difference-image priors to couple pairs of datasets within a longitudinal series. This approach is related to a variety of existing methods, none of which were directly applicable to the PET case. For example, penalising aspects of difference images has been used previously in the prior image constrained

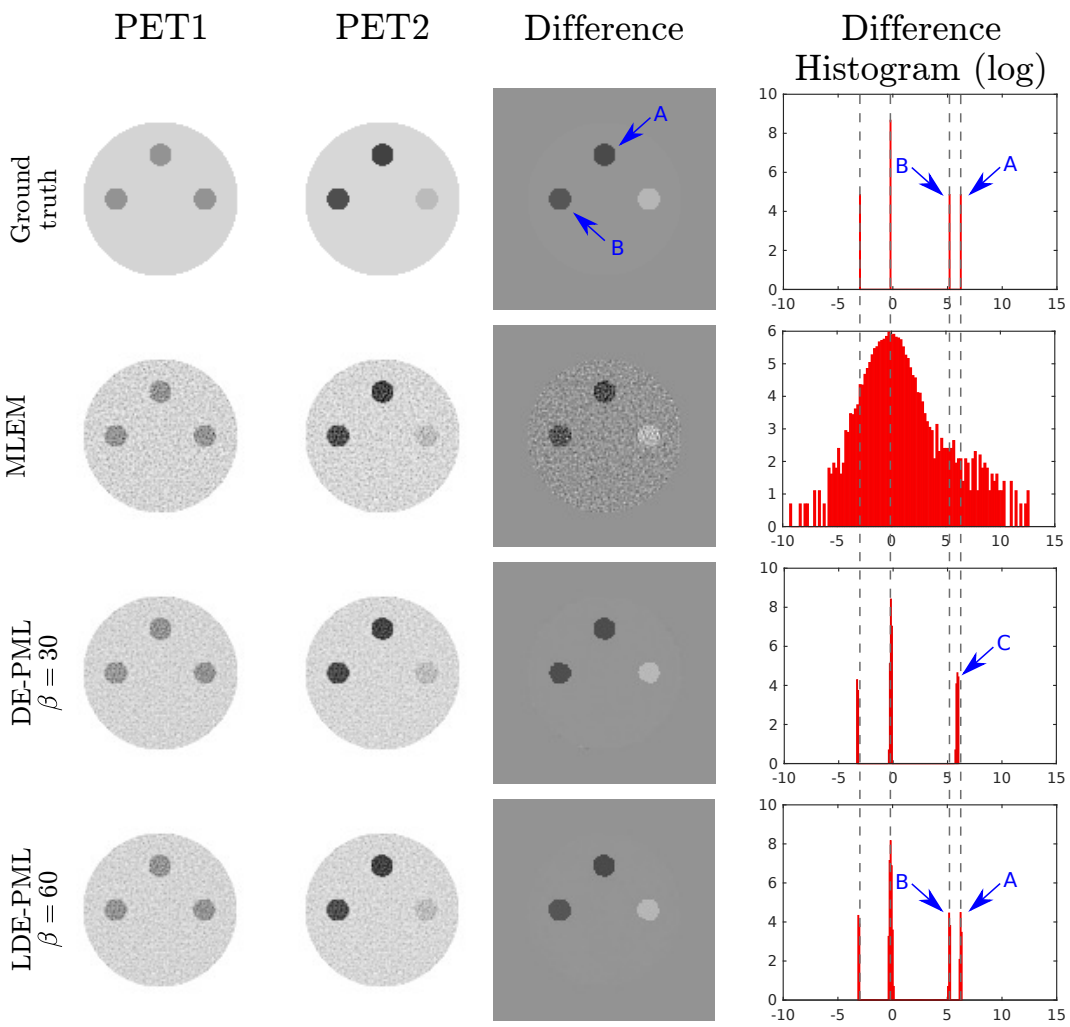


Figure 6.2: Simplified 2D example comparing the effects of a localised DE (LDE) and the DE priors. A pair of longitudinal ground truth images were produced which contained three regions of change (top row). Two of these regions were close in their difference values (A and B). When reconstructing simulated PET datasets with DE-PML (third row), the global nature of the method merged the peaks from A and B into a new peak (C), which corresponded to biased values for both regions A and B. When using the LDE prior with 15×15 neighbourhoods (bottom row), the peaks from regions A and B remained distinct, reducing the bias in regional quantification. Note that β values were tuned for both the DE-PML and LDE-PML methods to provide similar noise levels in the difference images.

reconstruction of multi-dataset CT [199–203] and MR [211,212] images, as described in Sections 2.4.2 and 2.4.3. In particular, both the sparsity and TV of difference images have been used, sometimes simultaneously [203]. However, in these cases the baseline image is pre-reconstructed and held constant during the reconstruction of the follow-up or difference images. This is because CT or MR images typically possess low image noise, reducing the risk of transferring noise between images. In fact, to minimise this risk even further, some of these multi-dataset CT image reconstruction methods include a TV penalty on the target image itself (c.f. Equation 2.49) [199]. The generally high level of intrinsic image noise in PET imaging led us to propose a simultaneous reconstruction approach. In this way, the proposed penalties encourage similar noise levels in all PET images using the entire PET data, thereby effectively sharing counts between the datasets.

The method of simultaneously reconstructing separate medical imaging datasets via the use of joint penalties has also been proposed in the past, most commonly for PET-MR image reconstruction [148, 220–224], as reviewed in Section 2.4.6. In this case, the types of prior that have been proposed have been informed by the complementary characteristics of the two imaging modalities. Because clinical MR images often exhibit good soft-tissue contrast and high spatial resolution, research has aimed to transfer MR edges into PET images (c.f. MR-guided PET image reconstruction, Section 2.4.1). Moreover, PET images have provided information to the MR images, potentially allowing increased MR data undersampling while avoiding the artefacts associated with such undersampling. Consequently, joint PET-MR reconstruction methods have typically sought to encourage common edges between the imaging modalities. In tumour monitoring PET studies however, neither image contains high-quality edge information, and tumour edges may change between scans. For these reasons, the proposed priors operated penalties in the difference image domain, using characteristics other than joint edges. It is interesting to note that simultaneous ET image reconstruction using joint edges has been performed before, in the context of dual-radiotracer rest/stress cardiac imaging using myocardial perfusion SPECT (Section 2.4.5) [215]. However, in

this case the images possessed high-contrast edges and the structure of interest (the left ventricular myocardium) had a relatively simple shape.

As well as improving the reconstructed PET images in a longitudinal series, the difference-image prior methods proposed in this thesis were observed to improve the difference images themselves. Difference images or change maps can be of interest in a range of ET contexts, for example in the use of ictal/interictal subtraction SPECT for the detection of epileptic foci [242–244] or the analysis of responding/non-responding voxels in oncology treatment response assessment [86, 88, 89]. We are aware of only two methods for improving difference images during image reconstruction in the ET literature [213, 214]. These two approaches reconstructed the difference image directly, using the tomographic forward model given by Equation 2.51. By the nature of the EM reconstruction algorithm used, these works reported improved visualisation of regions of change. Nonetheless, they did not report any improvement in image quality in terms of reconstruction error or quantification. In future it would be interesting to compare the difference-image prior methods proposed in this thesis to these direct difference image reconstruction methods. It is suspected that the extra assumptions included in the methods proposed in this thesis can further improve the image reconstruction performance relative to the direct reconstructions which include no prior knowledge or assumptions.

6.2 Multi-Tracer Guided PET Image Reconstruction

Chapter 5 introduced the concept of guiding the reconstruction of one PET radiotracer image with another, higher quality, PET image. Although guided image reconstruction has a long history in ET (see Section 2.4.1), to the best of the author's knowledge it has never been applied to multi-tracer PET studies. As noted above, dual-radiotracer SPECT images have been simultaneously reconstructed by encouraging joint edges [215], but in that case the image structure was simple and the motivation was the perfect alignment of both

scans.

In general, to perform guided image reconstruction, assumptions must be made about the correspondence of the images in question. In this thesis [^{18}F]-FDG and [^{11}C]-methionine PET images of the brain were considered, motivated by the observation of similar boundary locations (Figure 5.1). The biological argument for this is that non-specific radiotracers such as [^{18}F]-FDG and [^{11}C]-methionine will generally accumulate according to tissue boundaries, according to the characteristics of the tissues in question. Note that the assumption of similar boundaries made in the multi-tracer guidance of this thesis is weaker than the joint edges assumption of [215]. This is to account for a) the more structurally complex edges that occur in brain PET, where structure may be lost across scales close to the system resolution, b) the weaker edges present in the low-contrast [^{11}C]-methionine images, and c) the potential misalignment caused by repeated, as opposed to simultaneous, multi-tracer imaging. In a preliminary validation study using a coarse image grid, edge locations in the two radiotracers were shown to correspond down to a certain voxel size (Figure 5.16). Future validation making use of the same methodology with high-counts methionine images reconstructed into finer image grids would be required to conclude comprehensively the validity of guiding [^{11}C]-methionine image reconstruction with [^{18}F]-FDG images.

In order to prove that in principle there can be a benefit to transferring information from one PET image into the reconstruction of another, the weighted quadratic penalty (Equation 5.2) was used as the guidance method. Whereas the unweighted quadratic penalty seeks to penalise all pairwise differences within neighbourhoods, the inclusion of guidance weights allows the strength of this pairwise difference penalty to be modulated according to the information in the guiding image. A popular method of choosing these weights was proposed by Bowsher, where only the B most similar neighbours for each voxel have a weight of 1, and all other weights are set to 0 [168, 169]. In the case of anatomically guided ET image reconstruction using binary guidance weights, the accuracy of the alignment and the structural correspondence between the guiding image and the target ET image are of high importance.

To make the reconstructions more robust to these effects, Fessler proposed the use of blurred weights, representing the uncertainty in the location of an expected edge [135]. In the proposed multi-tracer method that uses a general non-binary measure of similarity between voxel pairs (Equation 5.3, similar to that used in [172]), the weights are naturally blurred by a combination of the hyperparameter values and the resolution of the guiding PET image. This results in an increased robustness against misalignment artefacts: the error in the registration of the attenuation maps may be less than the uncertainty suggested by the natural blur of the weights. In contrast to anatomically guided PET image reconstruction, which often aims to improve image resolution beyond the intrinsic PET resolution (i.e. perform partial volume correction), the proposed multi-tracer guided image reconstructions aimed to achieve noise reduction alone.

An important aspect of the calculation of the guidance weights is the sparsification of the weights used. If the similarity measure used were perfect, sparsification would not be necessary other than for computational reasons. However, similarity measures are imperfect, and so sparsification steps can improve the quality of reconstructed images (Figure 5.8). In this work a maximum value sparsification method was used, where only the n_{Max} maximum weights for each voxel were kept and all other weights set to zero. For computational efficiency, this was applied within neighbourhoods rather than globally. This approach is related to the sparsification used in the Bowsher method [169]. It is also known that performing maximum value sparsification using the similarity function in Equation 5.3 is equivalent to using a k-nearest neighbours sparsification (Equation 2.44a) of composite feature vectors containing the concatenation of patch intensities and location vectors [186]. By including the spatial information in the sparsification in this manner, the resulting sets of weights are more compact, which can aid the reconstruction of smaller structures that are unique to the target image [186]. Nonetheless, the maximum value sparsification method has the disadvantage of enforcing the use of n_{Max} similarities between voxel pairs in the reconstruction, which may be inappropriate for small features. Threshold-based sparsification methods

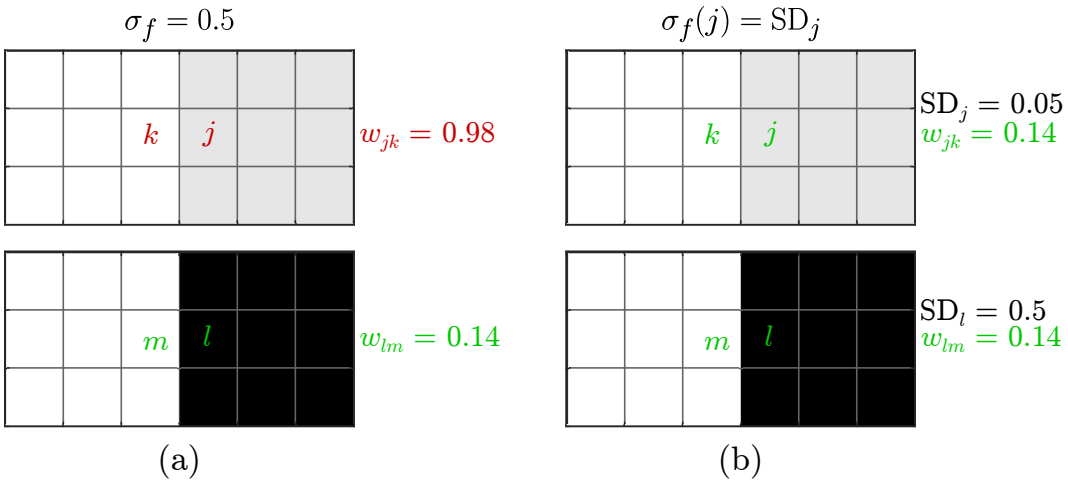


Figure 6.3: The principle of using a spatially variable σ_f value for calculation of guidance weights. (a) Use of a global σ_f value can lead to high similarity being observed over low-contrast edges (top), despite the clear existence of such edges. (b) By adapting the σ_f value according to the local image contrast (here using a standard deviation filter), similarity over low-contrast edges can be kept at levels similar to over the high contrast edge.

may alleviate this issue, allowing an adaptive number of non-zero weights per voxel. For the method proposed in Chapter 5, the weights could be sparsified by setting w_{jk} to zero for $w_{jk} < t$, where t is a threshold value between 0 and 1. Note that in the same way as maximum value sparsification of calculated weights is equivalent to applying k-nearest neighbours in feature space [186], thresholding calculated weights would be equivalent to using the ε-ball method (Equation 2.44c) proposed by Wang and Qi [180].

When calculating the guidance weights, the conventional approach is to use a global value of σ_f. The value of σ_f determines the strength of similarity provided over image edges. However, the optimum magnitude of σ_f depends on the magnitude of the edges being considered. Figure 6.3(a) shows how a global σ_f value can fail to reflect low-contrast edges in the calculated weights when the value is chosen based on high-contrast edges. By correcting local σ_f values, for example using the local standard deviation, the weights across the low-contrast edge can be made equivalent to those over the high-contrast edges (Figure 6.3(b)). Detailed investigation into the use of spatially-variant σ_f values remains for future work.

The weighted quadratic penalty is not the only method for incorporating

guidance information into a PET image reconstruction problem (see Section 2.4.1). Previous work has attempted to compare various guided reconstruction methods in the context of anatomically guided PET image reconstruction [147, 168, 172, 173], but it has proven difficult to fully characterise the wide range of available methods in terms of all relevant hyperparameters in an application-specific manner. Further investigation into the use of alternative methods will be necessary in the future in order to be able to make recommendations about which methods are most useful for the multi-tracer case. In particular, in the weighted quadratic penalty used in this thesis, the issue of possible cross-talk remains, for example in the blurring of small [^{11}C]-methionine specific features in Figure 5.12. One option to reduce the effects of such cross-talk is to include information from the target image in the calculation of the weights in an adaptive manner varying with iteration [170, 172]. Alternatively, using simultaneous reconstruction methods such as the DTV and DE priors proposed in this thesis for longitudinal oncology imaging could reduce cross-talk, and the use of joint edge priors like that used in [215] could be explored.

Interestingly, our preliminary results suggest that there can be a benefit to using multi-tracer guided PET image reconstruction compared to anatomically guided reconstruction alone. Figure 5.14 shows that [^{18}F]-FDG images can contain additional information compared to CT or MR images, due to the high contrast of [^{18}F]-FDG images compared to such anatomical images. Combining information from prior PET and anatomical images may enhance the guided reconstruction results. Ultimately, investigation of the optimum guidance images for PET image reconstruction needs to be addressed in an application-specific manner, where the range of PET and MR images acquired is the limiting factor in terms of how much prior information is available and how it can be used.

6.3 Comparison of Simultaneous and Guided Reconstruction Approaches

This thesis proposed simultaneous and guided image reconstruction methods, for the cases of oncology treatment monitoring and multi-tracer PET respectively. The methods were informed by the applications in question: by acting in the difference image domain the simultaneous reconstruction methods aimed to focus on the characteristic changes in longitudinal oncology scans, and in order to take advantage of the superior quality of some PET images, the guided reconstructions transferred spatial information from [^{18}F]-FDG images into the reconstruction of [^{11}C]-methionine images.

Of course, the choice between simultaneous and guided methods depends on the quality of the data at hand and the types of assumptions made. For example, simultaneous reconstructions were selected for the longitudinal oncology image reconstruction methods to avoid imprinting the noise present in the baseline image into the follow-up images. In principle though, if the baseline image were of a sufficiently high quality, the difference-image penalties used in this thesis could be applied to the difference between the target follow-up images and a fixed baseline image, similar to the PIRPLE method in CT image reconstruction [199, 200].

Similarly, although the [^{18}F]-FDG image was held fixed in the proposed multi-tracer image reconstructions, simultaneous methods would be applicable, similar to the cross-tracer priors proposed in myocardial SPECT imaging [215]. In fact, the proposed DTV and DE priors could be applicable to the multi-tracer case, given that they exhibit some level of contrast-invariance.

Importantly, the simultaneous difference-image reconstructions differ from the guided multi-tracer reconstructions in the amount and nature of noise reduction achieved. While the simultaneous methods seemed to posses an upper limit on noise reduction related to the number of counts in the entire data-series, the FDG-guided [^{11}C]-methionine reconstructions were able to achieve a much lower level of image noise. This is due to a difference in the domains in

which the priors denoise. The difference-image priors denoise in the difference image, which allows the individual PET images to exhibit higher levels of longitudinally correlated noise. In contrast, the weighted quadratic penalty used in the guided multi-tracer reconstructions operated in the [^{11}C]-methionine image directly, allowing a greater level of noise reduction to be achieved.

However, one important advantage of the difference-image prior methods as proposed in this thesis is that the noise texture obtained in the reconstructed images appears similar to that usually obtained by using higher counts levels (Figures 3.7). In this way, the difference-image priors achieve improvement to the PET images, but retain the characteristic texture of MLEM images (Figure 4.14). This more naturally-appearing noise reduction is attractive in terms of using these methods in routine clinical practice, avoiding the over-smoothed images often achieved by conventional penalised likelihood reconstruction methods.

6.4 Relation to Dynamic PET Image Reconstruction

The methods presented in this thesis are related to methods proposed in the context of dynamic PET image reconstruction, where the estimation of time-activity-curves or kinetic parameters is of interest. For example, guided reconstruction using the kernel method was originally proposed to allow enhanced reconstruction of early time frames in dynamic PET imaging [180]. Another example is the use of temporal priors to denoise reconstructed time-activity-curves, in a straightforward extension of spatial MRF priors (e.g. [245]). However, there are notable differences between dynamic reconstruction methods and multi-dataset reconstruction methods. Firstly, dynamic PET data are usually acquired continuously throughout the biological process of interest, which ensures that time-frames in the dynamic reconstruction are abutting. This means that it is easier to encourage temporal similarity between frames, since it is safe to assume that no large changes have occurred. In the longi-

tudinal case, the scans could quite feasibly be separated by months, which is ample time for very large changes to occur. Secondly, because the overall aim of dynamic reconstruction is often to estimate kinetic parameters according to pharmacokinetic compartmental models, there can be a greater level of robustness to errors in individual isolated voxel values throughout the dynamic series. On the other hand, in longitudinal studies the exact values of voxels and ROIs can be very important for a given single scan, often more so than the case of a single frame within a dynamic reconstruction. Finally, in dynamic PET contexts, the number of time frames is considerably higher than the typical number of scans available for longitudinal studies. This limitation of longitudinal studies means that it is more difficult to devise meaningful longitudinal models which would be suitable for direct incorporation into the image reconstruction method.

6.5 Experimental Limitations

In the simulation studies performed in the course of this thesis, some simplifications and assumptions were made. The PET data simulations used an analytic approach where Poisson noise was introduced to noise-free projection data of ground truth images, with effects like attenuation, randoms, and scatters modelled and applied in sinogram space. While Monte Carlo simulations allow more accurate data simulation since they model physical effects on an event-by-event basis, they are computationally expensive and are generally of most use when the effects of these physical processes are specifically of interest. The analytic approach taken in this thesis is conventionally used for the characterisation of PET image reconstruction methods [246], where the differences in reconstruction performance between methods is of interest rather than the absolute performance of each.

An important assumption made in the definition of the ground truth images in this thesis was that all tumours were homogeneous in their intensity. This assumption is common in the simulation of PET datasets, but in clinical PET imaging heterogeneous tumours are frequently encountered [241]. Test-

ing reconstruction methods on simulated heterogeneous tumours is difficult, since there is a general lack of models that allow many such heterogeneous tumours to be defined in a biologically meaningful way. The use of homogeneous tumours as used in this thesis allows for investigation of reconstruction performance as a function of tumour characteristics, permitting preliminary experiments in validating the proposed methods over a range of tumours (Figures 4.10 and 4.11). Of course, the relevance of tumour heterogeneity depends on the PET application being considered. For this reason, further validation should be performed on real datasets, where the natural variability of the disease being studied will be present. In such contexts, application-specific figures of merit would need to be used given the absence of ground truth images.

Throughout this thesis, measurement of image quality has primarily focused on reconstruction error (%RMSE or %MAE), or tumour mean values. Reconstruction error is a useful metric of reconstruction performance when a ground truth image is available, but it can be misleading in terms of clinical utility since it is unknown to what extent the ground truth images used in simulation studies reflect the spatial distributions of radiotracer uptake encountered *in vivo*. For this reason, the mean tumour intensity was included as a surrogate measure of the mean SUV that is typically used clinically for homogeneous tumours. The methods proposed in this thesis showed promise in simulation studies and real data case studies, but ultimately they will need to be validated on large numbers of PET datasets for specific contexts, for example head and neck cancer cases like that used in Chapter 4.

6.5.1 Algorithm Choice and Implementation

EM algorithms have been used exclusively in this thesis to solve the proposed optimisation problems. For the difference-image prior methods the one-step-late approximation first proposed by Green [143, 151] was implemented, which uses the gradient of the penalty function evaluated at the current image estimate to allow a closed form update similar to that of standard MLEM (Equation 2.33). The main disadvantage of this approach is that it is not guaranteed

to converge to a global maximum of the PML objective function and it can result in unstable reconstructions when the penalty strength β is large enough (c.f. the oscillatory behaviour seen in Figure 3.7, even where β was not excessive). However, the OSL method is simple to implement and can give valuable insight into the characteristics of selected priors. Furthermore, because the update formula is very similar to MLEM, the OSL method retains many of its benefits and converges at a similar rate (Figure 4.7), making comparisons to MLEM reconstructions fair.

In the multi-tracer guided PET image reconstruction work presented in Chapter 5, a modified EM algorithm was used: the method of De Pierro [153]. In this case, a surrogate for the quadratic penalty can be employed, allowing a separable update formula to be defined (see [154] for details). This method has the advantages of guaranteed convergence, simple implementation, and the ability to use any value for the penalty strength. In terms of comparison between PML reconstructions and MLEM reconstructions, the modified EM algorithm has a different convergence behaviour which hinders the comparison of both methods. In this thesis the standard approach of using a high number of iterations (200) was used to address this problem.

The modified EM algorithm can be applied to smooth ℓ_1-norm penalties [137, 142], making it potentially applicable to the DS and DTV priors. However, the convergence of the reconstruction is adversely affected, with the rate of convergence decreasing as the prior becomes increasingly non-smooth [137]. There are other algorithms that have been recently developed for the use of ℓ_1-norm priors in PET image reconstruction [137, 138, 140], but these methods are often less straightforward to implement, with elaborate bent-line-searches to be performed to avoid negative image values, or numerous hyperparameters to be tuned to obtain reliable performance. Investigation of applying these more advanced optimisation algorithms to the case of multi-dataset PET image reconstruction remains for future work.

6.5.2 The Role of Accurate Alignment Operators

Multi-dataset PET images are often misaligned when viewed in their native spaces, due to varying bed position and changes in subject pose between the scans, amongst others. For this reason, alignment of the image spaces is vital to permit the transfer of information between scans while ensuring good correspondence. In this thesis the existence of additional imaging data is assumed, from which alignment operators can be calculated. In general, for PET imaging, this is a reasonable assumption given that attenuation maps are routinely acquired with PET data to allow attenuation correction. Furthermore, these attenuation maps are typically acquired with CT or MR, exhibiting superior image resolution compared to the PET images to be aligned, thereby reducing potential misregistration errors. Nonetheless, the use of additional images to align two PET images can fail when there is motion between the acquisition of the attenuation maps and the PET data. For MR images in PET-MR protocols, this issue is mitigated by the fact that an attenuation map is generally acquired simultaneously to (a portion of) the PET data acquisition. For CT-based attenuation correction in PET-CT the risk of misalignment is greater due to the non-simultaneity of the data acquisitions.

In the experiments on difference-image priors for simultaneous reconstruction of longitudinal oncology PET images, non-rigid registration of co-acquired CT images was used to estimate the required alignment operators. Although this method produced good results, additional studies are required to verify the robustness of the method. Performing such a validation study would be vital in future, given recent research showing that non-rigid alignment of longitudinal oncology images can impact the calculation of tumour statistics [92]. To alleviate this problem, some research into voxel-wise analysis of longitudinal tumour changes have used rigid or affine alignment of tumours, forgoing accurate alignment in the remainder of the PET images [86, 89]. Ultimately, the alignment of PET images containing tumours remains an open question, although the methods proposed in Chapter 4 are general enough to include any alignment operator that can be expressed in a matrix form.

For the FDG-guided [^{11}C]-methionine image reconstruction experiments, a rigid alignment based on the MR-derived attenuation maps was considered suitable, because the brain does not typically undergo non-rigid motion, especially over the short time intervals typical in multi-tracer PET studies (<24 hr). The results of the real data test supported this assumption, with no visible misalignment errors.

As well as using side information to estimate the alignment operators, there is also the opportunity to estimate them from the PET data directly. Simultaneous estimation of images and alignment operators has been performed before in PET image reconstruction [230, 231, 233] and in longitudinal CT image reconstruction [199, 200]. These methods could be included in the techniques presented in this thesis, possibly as a way to allow the PET data to override or correct, where necessary, erroneous alignment estimated from the attenuation maps.

6.6 Future Prospects

Although this thesis focused on tumour response imaging and multi-tracer PET, the methods presented could be useful in a range of other applications. For example, the difference-image priors could help in the aforementioned reconstruction of ictal/inter-ictal SPECT scans for the detection of epileptic foci [242–244], the longitudinal observation of neurodegenerative diseases [27, 28], or brain activation and receptor mapping studies [50–55]. Ultimately, we hope that the work presented in this thesis will inspire the use of multi-dataset coupling of PET images during reconstruction in a range of contexts.

Nonetheless, there also remains much work to be done in fully characterising the benefits of the proposed methods in the applications considered within this work. The difference-image prior methods, for example, improve the visualisation of difference images, potentially allowing easier discrimination of regions of change from the background. In oncology treatment monitoring scans this may allow improved detection, localisation, or delineation of regions

of change. Future work will include investigating difference-image priors in terms of their receiver operating characteristics in order to determine in which cases these methods offer the most benefits.

As with many noise reduction methods in PET, there is also the opportunity to reduce patients' exposure to radiation by lowering the amount of injected radioactivity. For the difference-image priors, the maximum level of noise reduction possible is related to the number of scans in the longitudinal series. By coupling S datasets in the reconstruction it was observed that it is possible to reduce noise to levels similar to those typically observed when reconstructing datasets containing S times the number of counts, while providing the naturally-appearing noise reduction described previously. A corollary of this is that if there is a fixed amount of activity, A (in Bq), to be applied throughout a longitudinal series, it is possible to split this in to S scans of A/S Bq while maintaining similar noise levels equivalent to a single A Bq scan. Future work will investigate this splitting of the data, focusing on whether it is preferable to split the data evenly or unevenly between S scans, and on determining the maximum number of such splittings that is useful.

Chapter 7:
Conclusions

7.1 Contributions of this Thesis

The primary aim of this thesis was to propose and explore methods for sharing information between multiple PET datasets during image reconstruction in order to improve the output images. This was motivated by the existence of intra-subject similarity in such data-series.

With the exception of a handful of papers in the ET image reconstruction literature [213–215], this is an area that has historically been overlooked. This thesis presented two approaches for such multi-dataset PET image reconstruction: the application of difference-image priors in a simultaneous image reconstruction framework, and the use of high quality PET images to guide the reconstruction of poorer quality ones.

Difference-image priors were applied to oncology treatment response PET imaging. Initial priors were designed to encourage sparse difference images between a pair of oncology treatment monitoring scans, based on the assumption that the majority of voxels in such scans exhibit the same metabolic activity longitudinally. These were named difference sparsity (DS) priors. A 2D simulation study and 3D split real data experiments showed that the use of DS priors reduced image-wide reconstruction error and increased the contrast-to-noise ratio of tumours in reconstructed images. Following this, the simultaneous longitudinal image reconstruction framework was extended to allow the reconstruction of an arbitrary number of scans and to account for factors such

as scan misalignment and variations in injected activity levels. Furthermore, additional difference-image priors were defined, encouraging low entropy and low total variation in difference images (known as the DE and DTV priors respectively). Further 2D simulation studies demonstrated that when reconstructing S datasets, all three priors could reduce background image noise to levels approaching standard reconstruction methods applied to datasets with S times the number of counts. However, the three priors differed in their performance in terms of the representation of tumours in the reconstructed images. Of the three priors, the DE prior best preserved the clinically relevant tumour mean measure, while the DTV prior provided the lowest tumour reconstruction error levels due to high levels of noise reduction within the tumour itself. Application to a real data case study showed that the DTV prior was most robust to the errors associated with real data, such as the existence of regions of change not attributable to the tumour. Based on these results, recommendations for future difference-image priors were made, with a local DE prior showing improvements over the DE prior in a simplified 2D example. Ultimately, the use of difference-image priors for longitudinal oncology monitoring PET imaging has the potential to improve image quality via naturally appearing noise reduction, enhance the detection of regions of change, and reduce variability in the quantification of tumour activity.

Guided reconstruction was proposed for the case of multi-tracer PET imaging, with the aim of reducing image noise and improving quantification. A weighted quadratic penalty was adopted and applied to the case of [^{18}F]-FDG/[^{11}C]-methionine imaging of brain tumours. In this method, structural information was extracted from a fixed [^{18}F]-FDG image, and embedded in the reconstruction of a [^{11}C]-methionine dataset. 3D simulation studies explored the effects of the various hyperparameters of the method, before application to real [^{18}F]-FDG/[^{11}C]-methionine data. Guiding the reconstruction of the [^{11}C]-methionine dataset with [^{18}F]-FDG structural information reduced image noise levels while preserving edges. Furthermore, compared to CT- and MR- guided PET reconstructions, FDG-guidance produced lower levels of blurring in some areas of the [^{11}C]-methionine image due to the increased contrast

of the [^{18}F]-FDG image. General patterns of [^{11}C]-methionine uptake were visually unaffected by the use of FDG-guidance. Overall, the results of these experiments showed that multi-tracer guided reconstruction has the potential to improve the reconstruction of low-counts PET images when anatomical guidance information is not available, or is of insufficient quality.

7.2 Future Work

Following the promising results obtained within this thesis, there are many avenues of research for future work. The use of difference-image priors for the reconstruction of longitudinal PET images showed promise when applied to the challenging specific case of tumour monitoring. Future work involves improving upon the priors explored in this thesis (for example, using the LDE prior proposed in Chapter 6), including adaptation to heterogeneous tumours, and application-specific testing using clinical datasets. In addition, the difference-image priors are suitable for other contexts, such as ictal/inter-ictal SPECT imaging, longitudinal imaging of neurodegenerative diseases, or receptor mapping studies. Application of difference-image priors in these other contexts raises many opportunities for the improved detection and localisation of regions of change.

This thesis also proposed the use of guided reconstructions in multi-tracer PET. While we focused on [^{18}F]-FDG/[^{11}C]-methionine imaging, there are many other applications that may benefit from this reconstruction approach in areas such as oncology and neurology. Future work would include assessment of guided reconstructions in these other contexts in order to ascertain which benefit most from guidance. Furthermore, in this thesis a weighted quadratic penalty was used to apply the desired guidance, potentially leading to elevated reconstruction errors when there is a large mismatch between PET images. Future work would explore the applicability of other options available from the anatomically guided reconstruction literature, including the use of self-guidance information, that could further improve guided image reconstruction accuracy and robustness.

Publications

The work presented in this thesis is based on the following publications:

- S. Ellis and A. J. Reader, "Simultaneous maximum *a posteriori* longitudinal PET image reconstruction," *Phys. Med. Biol.*, vol. 62, no. 17, pp. 6963-6979, 2017.

- S. Ellis and A. J. Reader, "Penalized maximum likelihood simultaneous longitudinal PET image reconstruction with difference-image priors," *Med. Phys.*, vol. 45, no. 7, pp. 3001-3018, 2018.

- S. Ellis, A. Mallia, C. J. McGinnity, G. J. R. Cook, and A. J. Reader, "Multi-Tracer Guided PET Image Reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol.2, no. 5, pp. 499-509, 2018.

During the course of this PhD, the following works have been presented at international conferences:

- S. Ellis and A. J. Reader, "Synergistic Longitudinal PET Image Reconstruction," *2016 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).*, Strasbourg, France. 2016.

- S. Ellis and A. J. Reader, "Kernelised EM Image Reconstruction for Dual-Dataset PET Studies," *2016 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).*, Strasbourg, France. 2016.

- S. Ellis and A. J. Reader, "Longitudinal Multi-Dataset PET Image Reconstruction," *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).*, Atlanta, USA. 2017.

- S. Ellis, A. Mallia, C. J. McGinnity, G. J. R. Cook, and A. J. Reader, "Guided Image Reconstruction for Multi-Tracer PET," *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).*, Atlanta, USA. 2017.

Bibliography

[1] V. J. Cunningham, V. W. Pike, D. Bailey *et al.*, "A method of studying pharmacokinetics in man at picomolar drug concentrations," *Br. J. Clin. Pharmacol.*, vol. 32, no. 2, pp. 167–172, 1991.

[2] P. E. Kinahan and J. W. Fletcher, "Positron Emission Tomography-Computed Tomography Standardized Uptake Values in Clinical Practice and Assessing Response to Therapy," *Semin. Ultrasound CT MRI*, vol. 31, no. 6, pp. 496–505, 2010.

[3] R. J. Gillies, I. Robey, and R. A. Gatenby, "Causes and Consequences of Increased Glucose Metabolism of Cancers," *J. Nucl. Med.*, vol. 49, no. 6 (Suppl), pp. 24S–42S, 2008.

[4] A. F. Scarsbrook and S. F. Barrington, "PET-CT in the UK: current status and future directions," *Clin. Radiol.*, vol. 71, no. 7, pp. 673–690, 2016.

[5] E. Safaie, R. Matthews, and R. Bergamaschi, "PET scan findings can be false positive," *Tech. Coloproctol.*, vol. 19, no. 6, pp. 329–330, 2015.

[6] N. M. Long and C. S. Smith, "Causes and imaging features of false positives and false negatives on 18F-PET/CT in oncologic imaging," *Insights Imaging*, vol. 2, no. 6, pp. 679–698, 2011.

[7] P. Raj, L. Kaman, R. Singh *et al.*, "Sensitivity and specificity of FDG PET-CT scan in detecting lymph node metastasis in operable periampullary tumours in correlation with the final histopathology after curative surgery," *Updates Surg.*, vol. 65, no. 2, pp. 103–107, 2013.

[8] H. J. Adams and T. C. Kwee, "Proportion of false-positive lesions at interim and end-of-treatment FDG-PET in lymphoma as determined by histology: Systematic review and meta-analysis," *Eur. J. Radiol.*, vol. 85, no. 11, pp. 1963–1970, 2016.

[9] H. W. D. Yeung, R. K. Grewal, M. Gonen *et al.*, "Patterns of 18F-FDG Uptake in Adipose Tissue and Muscle: A Potential Source of False-Positives for PET," *J. Nucl. Med.*, vol. 44, no. 11, pp. 1789–96, 2003.

[10] S. J. Rosenbaum, T. Lind, G. Antoch, and A. Bockisch, "False-Positive FDG PET Uptake - the Role of PET/CT," *Eur. Radiol.*, vol. 16, no. 5, pp. 1054–1065, 2006.

[11] C. Hofer, C. Laubenbacher, T. Block *et al.*, "Fluorine-18-Fluorodeoxyglucose Positron Emission Tomography Is Useless for the Detection of Local Recurrence after Radical Prostatectomy," *Eur. Urol.*, vol. 36, no. 1, pp. 31–35, 1999.

[12] R. J. Hicks, "Beyond FDG: novel PET tracers for cancer imaging," *Cancer Imaging*, vol. 4, no. 1, pp. 22–24, 2003.

[13] W. A. Weber, "Positron Emission Tomography As an Imaging Biomarker," *J. Clin. Oncol.*, vol. 24, no. 20, pp. 3282–3292, 2006.

[14] M. A. Pantaleo, M. Nannini, A. Maleddu *et al.*, "Conventional and novel PET tracers for imaging in oncology in the era of molecular therapy," *Cancer Treat. Rev.*, vol. 34, no. 2, pp. 103–121, 2008.

[15] K. K. S. Sai, Z. Zachar, P. M. Bingham, and A. Mintz, "Metabolic PET Imaging in Oncology," *Am. J. Roentgenol.*, vol. 209, no. 2, pp. 270–276, 2017.

[16] P. Bustany, M. Chatel, J. M. Derlon *et al.*, "Brain tumor protein synthesis and histological grades: A study by positron emission tomography (PET) with C11-L-Methionine," *J. Neurooncol.*, vol. 3, no. 4, pp. 397–404, 1986.

[17] G. J. R. Cook, M. N. Maisey, and I. Fogelman, "Normal variants, artefacts and interpretative pitfalls in PET imaging with 18-fluoro-2-deoxyglucose and carbon-11 methionine," *Eur. J. Nucl. Med.*, vol. 26, no. 10, pp. 1363–1378, 1999.

[18] R. Nuñez, H. A. Macapinlac, H. W. D. Yeung *et al.*, "Combined 18F-FDG and 11C-Methionine PET Scans in Patients with Newly Progressive Metastatic Prostate Cancer," *J. Nucl. Med.*, vol. 43, no. 1, pp. 46–55, 2002.

[19] A. F. Shields, "PET Imaging with 18F-FLT and Thymidine Analogs: Promise and Pitfalls," *J. Nucl. Med.*, vol. 44, no. 9, pp. 1432–4, 2003.

[20] J. C. Seung, S. K. Jae, H. K. Jeong *et al.*, "[18F]3'-deoxy-3'-fluorothymidine PET for the diagnosis and grading of brain tumors," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 32, no. 6, pp. 653–659, 2005.

[21] H. Wester, M. Herz, W. Weber *et al.*, "Synthesis and Radiopharmacology of O-(2- fluoroethyl)-L-Tyrosine for Tumor Imaging," *J. Nucl. Med.*, vol. 40, no. 1, pp. 205–12, 1999.

[22] D. Pauleit, F. Floeth, H. Herzog *et al.*, "Whole-body distribution and dosimetry of O-(2-[18F]fluoroethyl)-L-tyrosine," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 30, no. 4, pp. 519–524, 2003.

[23] V. Dunet, A. Pomoni, A. Hottinger *et al.*, "Performance of 18F-FET versus 18F-FDG-PET for the diagnosis and grading of brain tumors: systematic review and meta-analysis," *Neuro. Oncol.*, vol. 18, no. 3, pp. 426–434, 2016.

[24] R. K. Kulshrestha, S. Vinjamuri, A. England *et al.*, "The Role of 18F-Sodium Fluoride PET/CT Bone Scans in the Diagnosis of Metastatic Bone Disease from Breast and Prostate Cancer," *J. Nucl. Med. Technol.*, vol. 44, no. 4, pp. 217–222, 2016.

[25] S. Schwarzenboeck, M. Souvatzoglou, and B. J. Krause, "Choline PET and PET/CT in Primary Diagnosis and Staging of Prostate Cancer," *Theranostics*, vol. 2, no. 3, pp. 318–330, 2012.

[26] L. Mosconi, "Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 32, no. 4, pp. 486–510, 2005.

[27] L. Mosconi, R. Mistur, R. Switalski *et al.*, "FDG-PET changes in brain glucose metabolism from normal cognition to pathologically verified Alzheimer's disease," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 36, no. 5, pp. 811–822, 2009.

[28] W. J. Jagust, D. Bandy, K. Chen *et al.*, "The Alzheimer's Disease Neuroimaging Initiative positron emission tomography core," *Alzheimer's Dement.*, vol. 6, no. 3, pp. 221–229, 2010.

[29] H. Oakley, S. L. Cole, S. Logan *et al.*, "Intraneuronal β-Amyloid Aggregates, Neurodegeneration, and Neuron Loss in Transgenic Mice with Five Familial Alzheimers Disease Mutations: Potential Factors in Amyloid Plaque Formation," *J. Neurosci.*, vol. 26, no. 40, pp. 10 129–10 140, 2006.

[30] C. R. Jack, D. S. Knopman, W. J. Jagust *et al.*, "Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade," *Lancet Neurol.*, vol. 9, no. 1, pp. 119–128, 2010.

[31] W. E. Klunk, H. Engler, A. Nordberg *et al.*, "Imaging Brain Amyloid in Alzheimer's Disease with Pittsburgh Compound-B," *Ann. Neurol.*, vol. 55, no. 3, pp. 306–319, 2004.

[32] J. C. Price, W. E. Klunk, B. J. Lopresti *et al.*, "Kinetic modeling of amyloid binding in humans using PET imaging and Pittsburgh Compound-B," *J. Cereb. Blood Flow Metab.*, vol. 25, no. 11, pp. 1528–1547, 2005.

[33] D. F. Wong, P. B. Rosenberg, Y. Zhou *et al.*, "In Vivo Imaging of Amyloid Deposition in Alzheimer Disease Using the Radioligand 18F-AV-45 (Flobetapir F 18)," *J. Nucl. Med.*, vol. 51, no. 6, pp. 913–920, 2010.

[34] H. Barthel, H.-J. Gertz, S. Dresel *et al.*, "Cerebral amyloid-β PET with florbetaben (18F) in patients with Alzheimer's disease and healthy controls: a multicentre phase 2 diagnostic study," *Lancet Neurol.*, vol. 10, no. 5, pp. 424–435, 2011.

[35] A. Wimo, B. Winblad, and L. Jönsson, "An estimate of the total worldwide societal costs of dementia in 2005," *Alzheimer's Dement.*, vol. 3, no. 2, pp. 81–91, 2007.

[36] N. D. Volkow, J. S. Fowler, S. J. Gatley *et al.*, "PET Evaluation of the Dopamine System of the Human Brain," *J. Nucl. Med.*, vol. 37, no. 7, pp. 1242–1256, 1996.

[37] P. Cumming, "Absolute Abundances and Affinity States of Dopamine Receptors in Mammalian Brain: A Review," *Synapse*, vol. 65, no. 9, pp. 892–909, 2011.

[38] R. D. Badgaiyan, "Imaging dopamine neurotransmission in live human brain," *Prog. Brain Res.*, vol. 211, pp. 165–182, 2014.

[39] W.-D. Heiss and K. Herholz, "Brain Receptor Imaging," *J. Nucl. Med.*, vol. 47, no. 2, pp. 302–312, 2006.

[40] S. Seo, S. J. Kim, D. S. Lee, and J. S. Lee, "Recent advances in parametric neuroreceptor mapping with dynamic PET: basic concepts and graphical analyses," *Neurosci. Bull.*, vol. 30, no. 5, pp. 733–754, 2014.

[41] T. H. Schindler, H. R. Schelbert, A. Quercioli, and V. Dilsizian, "Cardiac PET Imaging for the Detection and Monitoring of Coronary Artery Disease and Microvascular Health," *JACC Cardiovasc. Imaging*, vol. 3, no. 6, pp. 623–640, 2010.

[42] J. Machac, "Cardiac Positron Emission Tomography Imaging," *Semin. Nucl. Med.*, vol. 35, no. 1, pp. 17–36, 2005.

[43] M. F. Di Carli, S. Dorbala, J. Meserve *et al.*, "Clinical Myocardial Perfusion PET/CT," *J. Nucl. Med.*, vol. 48, no. 5, pp. 783–793, 2007.

[44] M. R. Dweck, M. W. Chow, N. V. Joshi *et al.*, "Coronary Arterial 18F-Sodium Fluoride Uptake: A Novel Marker of Plaque Biology," *J. Am. Coll. Cardiol.*, vol. 59, no. 17, pp. 1539–1548, 2012.

[45] W. Chen and V. Dilsizian, "Targeted PET/CT Imaging of Vulnerable Atherosclerotic Plaques: Microcalcification with Sodium Fluoride and Inflammation with Fluorodeoxyglucose," *Curr. Cardiol. Rep.*, vol. 15, no. 6, p. 364, 2013.

[46] N. V. Joshi, A. T. Vesey, M. C. Williams *et al.*, "18F-fluoride positron emission tomography for identification of ruptured and high-risk coronary atherosclerotic plaques: a prospective clinical trial," *Lancet*, vol. 383, pp. 705–713, 2014.

[47] C. R. Jack, V. J. Lowe, S. D. Weigand *et al.*, "Serial PIB and MRI in normal, mild cognitive impairment and Alzheimers disease: implications for sequence of pathological events in Alzheimers disease," *Brain*, vol. 132, no. 5, pp. 1355–1365, 2009.

[48] H. Engler, A. Forsberg, O. Almkvist *et al.*, "Two-year follow-up of amyloid deposition in patients with Alzheimer's disease," *Brain*, vol. 129, no. 11, pp. 2856–2866, 2006.

[49] J. Sevigny, P. Chiao, T. Bussière *et al.*, "The antibody aducanumab reduces Aβ plaques in Alzheimer's disease," *Nature*, vol. 537, pp. 50–56, 2016.

[50] M. Corbetta, F. M. Miezin, G. L. Shulman, and S. E. Petersen, "A PET Study of Visuospatial Attention," *J. Neurosci.*, vol. 13, no. 3, pp. 1202–1226, 1993.

[51] A. Egerton, M. A. Mehta, A. J. Montgomery *et al.*, "The dopaminergic basis of human behaviors: A review of molecular imaging studies," *Neurosci. Biobehav. Rev.*, vol. 33, no. 7, pp. 1109–1132, 2009.

[52] D. J. Scott, C. S. Stohler, R. A. Koeppe, and J. K. Zubieta, "Time-Course of Change in [11C]Carfentanil and [11C]Raclopride Binding Potential After a Nonpharmacological Challenge," *Synapse*, vol. 61, no. 9, pp. 707–714, 2007.

[53] V. N. Salimpoor, M. Benovoy, K. Larcher *et al.*, "Anatomically distinct dopamine release during anticipation and experience of peak emotion to music," *Nat. Neurosci.*, vol. 14, no. 2, pp. 257–262, 2011.

[54] R. E. Carson, A. Breier, A. De Bartolomeis *et al.*, "Quantification of Amphetamine-Induced Changes in [11C]Raclopride Binding with Continuous Infusion," *J. Cereb. Blood Flow Metab.*, vol. 17, no. 4, pp. 437–447, 1997.

[55] A. Breier, T.-P. Su, R. Saunders *et al.*, "Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations: Evidence from a novel positron emission tomography method," *Proc. Natl. Acad. Sci.*, vol. 94, no. 6, pp. 2569–2574, 1997.

[56] E. D. Morris and K. K. Yoder, "Positron emission tomography displacement sensitivity: predicting binding potential change for positron emission tomography tracers based on their kinetic characteristics," *J. Cereb. Blood Flow Metab.*, vol. 27, no. 3, pp. 606–617, 2007.

[57] E. D. Morris, C. C. Constantinescu, J. M. Sullivan *et al.*, "Noninvasive visualization of human dopamine dynamics from PET images," *Neuroimage*, vol. 51, no. 1, pp. 135–144, 2010.

[58] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. Frackowiak, "Comparing Functional (PET) Images: The Assessment of Significant Change," *J. Cereb. Blood Flow Metab.*, vol. 11, no. 4, pp. 690–699, 1991.

[59] B. J. Chow, K. Ananthasubramaniam, R. A. DeKemp *et al.*, "Comparison of treadmill exercise versus dipyridamole stress with myocardial perfusion imaging using rubidium-82 positron emission tomography," *J. Am. Coll. Cardiol.*, vol. 45, no. 8, pp. 1227–1234, 2005.

[60] B. J. Chow, R. S. Beanlands, A. Lee *et al.*, "Treadmill Exercise Produces Larger Perfusion Defects Than Dipyridamole Stress N-13 Ammonia Positron Emission Tomography," *J. Am. Coll. Cardiol.*, vol. 47, no. 2, pp. 411–416, 2006.

[61] T. M. Bateman, "Cardiac Positron Emission Tomography and the Role of Adenosine Pharmacologic Stress," in *Am. J. Cardiol.*, vol. 94, no. 2A. Elsevier, 2004, pp. 19–24.

[62] A. Zhu, D. Lee, and H. Shim, "Metabolic Positron Emission Tomography Imaging in Cancer Detection and Therapy Response," *Semin. Oncol.*, vol. 38, no. 1, pp. 55–69, 2011.

[63] P. Therasse, S. G. Arbuck, E. A. Eisenhauer *et al.*, "New Guidelines to Evaluate the Response to Treatment in Solid Tumors," *J. Natl. Cancer Inst.*, vol. 92, no. 3, pp. 205–216, 2000.

[64] R. L. Wahl, H. Jacene, Y. Kasamon, and M. A. Lodge, "From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumors," *J. Nucl. Med.*, vol. 50, no. 5 (Suppl), pp. 122S–150S, 2009.

[65] N. Avril, S. Sassen, B. Schmalfeldt *et al.*, "Prediction of response to neoadjuvant chemotherapy by sequential F-18-fluorodeoxyglucose positron emission tomography in patients with advanced-stage ovarian cancer," *J. Clin. Oncol.*, vol. 23, no. 30, pp. 7445–7453, 2005.

[66] P. Cheebsumon, L. M. Velasquez, C. J. Hoekstra *et al.*, "Measuring response to therapy using FDG PET: semi-quantitative and full kinetic analysis," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 38, no. 5, pp. 832–842, 2011.

[67] T. Kishino, H. Hoshikawa, Y. Nishiyama *et al.*, "Usefulness of 3'-Deoxy-3'-18F-Fluorothymidine PET for Predicting Early Response to Chemoradiotherapy in Head and Neck Cancer," *J. Nucl. Med.*, vol. 53, no. 10, pp. 1–7, 2012.

[68] F. Lordick, K. Ott, B.-J. Krause *et al.*, "PET to assess early metabolic response and to guide treatment of adenocarcinoma of the oesophagogastric junction: the MUNICON phase II trial," *Lancet Oncol.*, vol. 8, no. 9, pp. 797–805, 2007.

[69] P. Georg, P. Andrzejewski, P. Baltzer *et al.*, "Changes in Tumor Biology During Chemoradiation of Cervix Cancer Assessed by Multiparametric MRI and Hypoxia PET," *Mol. Imaging Biol.*, vol. 20, no. 1, pp. 160–169, 2018.

[70] S. Stroobants, J. Goeminne, M. Seegers *et al.*, "18FDG-Positron emission tomography for the early prediction of response in advanced soft tissue sarcoma treated with imatinib mesylate (Glivec®)," *Eur. J. Cancer*, vol. 39, no. 14, pp. 2012–2020, 2003.

[71] W. A. Weber, K. Ott, K. Becker *et al.*, "Prediction of Response to Preoperative Chemotherapy in Adenocarcinomas of the Esophagogastric Junction by Metabolic Imaging," *J. Clin. Oncol.*, vol. 19, no. 12, pp. 3058–3065, 2001.

[72] L. Kenny, R. C. Coombes, D. M. Vigushin *et al.*, "Imaging early changes in proliferation at 1 week post chemotherapy: a pilot study in breast cancer patients with 3'-deoxy-3'-[18F]fluorothymidine positron emission tomography," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 34, no. 9, pp. 1339–1347, 2007.

[73] T. Ishimori, T. Saga, Y. Nagata *et al.*, "18F-FDG and 11C-methionine PET for evaluation of treatment response of lung cancer after stereotactic radiotherapy," *Ann. Nucl. Med.*, vol. 18, no. 8, pp. 669–674, 2004.

[74] L. Dercle, R.-D. Seban, J. Lazarovici *et al.*, "18F-FDG PET and CT Scans Detect New Imaging Patterns of Response and Progression in Patients with Hodgkin Lymphoma Treated by Anti-Programmed Death 1 Immune Checkpoint Inhibitor," *J. Nucl. Med.*, vol. 59, no. 1, pp. 15–24, 2018.

[75] W. Chen, S. Delaloye, D. H. Silverman *et al.*, "Predicting Treatment Response of Malignant Gliomas to Bevacizumab and Irinotecan by Imaging Proliferation with [18F] Fluorothymidine Positron Emission Tomography: A Pilot Study," *J. Clin. Oncol.*, vol. 25, no. 30, pp. 4714–4721, 2007.

[76] J. Nuutinen, P. Sonninen, P. Lehikoinen *et al.*, "Radiotherapy treatment planning and long-term follow-up with [11C]methionine PET in patients with low-grade astrocytoma," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 48, no. 1, pp. 43–52, 2000.

[77] J. Schwarzenberg, J. Czernin, T. F. Cloughesy *et al.*, "3'-Deoxy-3'-18F-Fluorothymidine PET and MRI for Early Survival Predictions in Patients with Recurrent Malignant Glioma Treated with Bevacizumab," *J. Nucl. Med.*, vol. 53, no. 1, pp. 29–36, 2012.

[78] B. S. Pio, C. K. Park, R. Pietras *et al.*, "Usefulness of 3'-[F-18]Fluoro-3'-deoxythymidine with Positron Emission Tomography in Predicting Breast Cancer Response to Therapy," *Mol. Imaging Biol.*, vol. 8, no. 1, pp. 36–42, 2006.

[79] M. A. Rueger, M. Ameli, H. Li *et al.*, "[18F]FLT PET for Non-Invasive Monitoring of Early Response to Gene Therapy in Experimental Gliomas," *Mol. Imaging Biol.*, vol. 13, no. 3, pp. 547–557, 2011.

[80] M. Schelling, N. Avril, J. Naehrig *et al.*, "Positron Emission Tomography Using [18F]Fluorodeoxyglucose for Monitoring Primary Chemotherapy in Breast Cancer," *J. Clin. Oncol.*, vol. 18, no. 8, pp. 1689–1695, 2000.

[81] M. Vanderhoek, S. B. Perlman, and R. Jeraj, "Impact of Different Standardized Uptake Value Measures on PET-Based Quantification of Treatment Response," *J. Nucl. Med.*, vol. 54, no. 8, pp. 1188–1194, 2013.

[82] E. Brun, E. Kjellén, J. Tennvall *et al.*, "FDG PET studies during treatment: Prediction of therapy outcome in head and neck squamous cell carcinoma," *Head Neck*, vol. 24, no. 2, pp. 127–135, 2002.

[83] W. A. Weber, V. Petersen, B. Schmidt *et al.*, "Positron Emission Tomography in Non-Small-Cell Lung Cancer: Prediction of Response to Chemotherapy by Quantitative Assessment of Glucose Use," *J. Clin. Oncol.*, vol. 21, no. 14, pp. 2651–2657, 2003.

[84] A. F. Shields, J. R. Grierson, B. M. Dohmen *et al.*, "Imaging proliferation in vivo with [F-18]FLT and positron emission tomography," *Nat. Med.*, vol. 4, no. 11, pp. 1334–1336, 1998.

[85] V. A. Holthoff, K. Herholz, F. Berthold *et al.*, "In Vivo Metabolism of Childhood Posterior Fossa Tumors and Primitive Neuroectodermal Tumors before and after Treatment," *Cancer*, vol. 72, no. 4, pp. 1394–1403, 1993.

[86] H. Necib, C. Garcia, A. Wagner *et al.*, "Detection and Characterization of Tumor Changes in 18F-FDG PET Patient Monitoring Using Parametric Imaging," *J. Nucl. Med.*, vol. 52, no. 3, pp. 354–361, 2011.

[87] S. David, D. Visvikis, C. Roux, and M. Hatt, "Multi-observation PET image analysis for patient follow-up quantitation and therapy assessment," *Phys. Med. Biol.*, vol. 56, no. 18, pp. 5771–5788, 2011.

[88] E. Schreibmann, A. F. Waller, I. Crocker *et al.*, "Voxel clustering for quantifying PET-based treatment response assessment," *Med. Phys.*, vol. 40, no. 1, p. 012401, 2013.

[89] P. Vera, B. Dubray, O. Palie *et al.*, "Monitoring tumour response during chemo-radiotherapy: a parametric method using FDG-PET/CT images

in patients with oesophageal cancer," *EJNMMI Res.*, vol. 4, no. 1, p. 12, 2014.

[90] M. Wang, N. Guo, G. Hu *et al.*, "A novel approach to assess the treatment response using Gaussian random field in PET," *Med. Phys.*, vol. 43, no. 2, pp. 833–842, 2016.

[91] S. Tan, H. Zhang, Y. Zhang *et al.*, "Predicting pathologic tumor response to chemoradiotherapy with histogram distances characterizing longitudinal changes in 18F-FDG uptake patterns," *Med. Phys.*, vol. 40, no. 10, p. 101707, 2013.

[92] G. S. M. A. Kerner, A. Fischer, M. J. B. Koole *et al.*, "Evaluation of elastix-based propagated align algorithm for VOI- and voxel-based analysis of longitudinal 18F-FDG PET/CT data from patients with non-small cell lung cancer (NSCLC)," *EJNMMI Res.*, vol. 5, no. 1, p. 15, 2015.

[93] A. Andreyev and A. Celler, "Dual-isotope PET using positron-gamma emitters," *Phys. Med. Biol.*, vol. 56, no. 14, pp. 4539–4556, 2011.

[94] D. J. Kadrmas, T. C. Rust, and J. M. Hoffman, "Single-scan dual-tracer FLT+FDG PET tumor characterization," *Phys. Med. Biol.*, vol. 58, no. 3, pp. 429–449, 2013.

[95] J. Verhaeghe and A. J. Reader, "Simultaneous water activation and glucose metabolic rate imaging with PET," *Phys. Med. Biol.*, vol. 58, no. 3, pp. 393–411, 2013.

[96] G. E. Fakhri, C. M. Trott, A. Sitek *et al.*, "Dual-Tracer PET Using Generalized Factor Analysis of Dynamic Sequences," *Mol. Imaging Biol.*, vol. 15, no. 6, pp. 666–674, 2013.

[97] A. Andreyev, A. Sitek, and A. Celler, "EM reconstruction of dual isotope PET using staggered injections and prompt gamma positron emitters," *Med. Phys.*, vol. 41, no. 2, p. 022501, 2014.

[98] J.-K. Chung, Y. K. Kim, S. Kim *et al.*, "Usefulness of 11C-methionine PET in the evaluation of brain lesions that are hypo- or isometabolic on 18F-FDG PET," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 29, no. 2, pp. 176–182, 2002.

[99] H. Jadvar, "Prostate Cancer: PET with 18F-FDG, 18F- or 11C-Acetate, and 18F- or 11C-Choline," *J. Nucl. Med.*, vol. 52, no. 1, pp. 81–89, 2011.

[100] C.-L. Ho, S. C. H. Yu, and D. W. C. Yeung, "11C-Acetate PET Imaging in Hepatocellular Carcinoma and Other Liver Masses," *J. Nucl. Med.*, vol. 44, no. 2, pp. 213–221, 2003.

[101] D. Pellegrino, F. Cicchetti, X. Wang *et al.*, "Modulation of Dopaminergic and Glutamatergic Brain Function: PET Studies on Parkinsonian Rats," *J. Nucl. Med.*, vol. 48, no. 7, pp. 1147–1153, 2007.

[102] O. Howes, R. McCutcheon, and J. Stone, "Glutamate and dopamine in schizophrenia: An update for the 21st century," *J. Psychopharmacol.*, vol. 29, no. 2, pp. 97–115, 2015.

[103] S. E. Derenzo, "Mathematical Removal of Positron Range Blurring in High Resolution Tomography," *IEEE Trans. Nucl. Sci.*, vol. 33, no. 1, pp. 565–569, 1986.

[104] S. R. Cherry, J. A. Sorenson, and M. E. Phelps, "Chapter 18 - Positron Emission Tomography," in *Physics in Nuclear Medicine*, 4th ed. Philadelphia: W.B. Saunders, 2012, pp. 307–343.

[105] S. Vandenberghe, E. Mikhaylova, E. D'Hoe *et al.*, "Recent developments in time-of-flight PET," *EJNMMI Phys.*, vol. 3, no. 1, p. 3, 2016.

[106] C. Burger, G. Goerres, S. Schoenes *et al.*, "PET attenuation coefficients from CT images: experimental evaluation of the transformation of CT into PET 511-keV attenuation coefficients," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 29, no. 7, pp. 922–927, 2002.

[107] H. Engel, H. Steinert, A. Buck *et al.*, "Whole-Body PET: Physiological and Artifactual Fluorodeoxyglucose Accumulations," *J. Nucl. Med.*, vol. 37, no. 3, pp. 441–446, 1996.

[108] S. R. Meikle, D. L. Bailey, P. K. Hooper *et al.*, "Simultaneous Emission and Transmission Measurements for Attenuation Correction in Whole-Body PET," *J. Nucl. Med.*, vol. 36, no. 9, pp. 1680–1688, 1995.

[109] R. J. Smith, J. S. Karp, G. Muehllehner *et al.*, "Singles Transmission Scans Performed Post-Injection for Quantitative Whole body PET imaging," *IEEE Trans. Nucl. Sci.*, vol. 44, no. 3, pp. 1329–1335, 1997.

[110] E. Kamel, T. F. Hany, C. Burger *et al.*, "CT vs 68Ge attenuation correction in a combined PET/CT system: evaluation of the effect of lowering the CT tube current," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 29, no. 3, pp. 346–350, 2002.

[111] P. E. Kinahan, D. W. Townsend, T. Beyer, and D. Sashin, "Attenuation correction for a combined 3D PET/CT scanner," *Med. Phys.*, vol. 25, no. 10, pp. 2046–2063, 1998.

[112] P. E. Kinahan, B. H. Hasegawa, and T. Beyer, "X-Ray-Based Attenuation Correction for Positron Emission Tomography/Computed Tomography Scanners," *Semin. Nucl. Med.*, vol. 33, no. 3, pp. 166–179, 2003.

[113] G. Wagenknecht, H.-J. Kaiser, F. M. Mottaghy, and H. Herzog, "MRI for attenuation correction in PET: methods and challenges," *Magn. Reson. Mater. Physics, Biol. Med.*, vol. 26, no. 1, pp. 99–113, 2013.

[114] G. Delso and J. Nuyts, "PET/MRI: Attenuation Correction," in *PET/MRI in Oncology: Current Clinical Applications*, A. Iagaru, T. Hope, and P. Veit-Haibach, Eds. Cham: Springer International Publishing, 2018, pp. 53–75.

[115] H. Zaidi and K. F. Koral, "Scatter modelling and compensation in emission tomography," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 31, no. 5, pp. 761–782, 2004.

[116] C. S. Levin and E. J. Hoffman, "Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution," *Phys. Med. Biol.*, vol. 44, no. 3, pp. 781–799, 1999.

[117] D. Yvon, J.-P. Renault, G. Tauzin *et al.*, "CaLIPSO: An Novel Detector Concept for PET Imaging," *IEEE Trans. Nucl. Sci.*, vol. 61, no. 1, pp. 60–66, 2014.

[118] A. Rahmim, J. Qi, and V. Sossi, "Resolution modeling in PET imaging: Theory, practice, benefits, and pitfalls," *Med. Phys.*, vol. 40, no. 6, p. 064301, 2013.

[119] L. Theodorakis, G. Loudos, V. Prassopoulos *et al.*, "A review of PET normalization: striving for count rate uniformity," *Nucl. Med. Commun.*, vol. 34, no. 11, pp. 1033–1045, 2013.

[120] M. A. Belzunce and A. J. Reader, "Time-invariant component-based normalization for a simultaneous PET-MR scanner," *Phys. Med. Biol.*, vol. 61, no. 9, pp. 3554–3571, 2016.

[121] B. M. Mazoyer, M. S. Roos, and R. H. Huesman, "Dead time correction and counting statistics for positron tomography," *Phys. Med. Biol.*, vol. 30, no. 5, pp. 385–399, 1985.

[122] M. E. Daube-Witherspoon and G. Muehllehner, "Treatment of axial data in three-dimensional PET." *J. Nucl. Med.*, vol. 28, no. 11, pp. 1717–1724, 1987.

[123] P. E. Kinahan and J. G. Rogers, "Analytic 3D Image Reconstruction Using All Detected Events," *IEEE Trans. Nucl. Sci.*, vol. 36, no. 1, pp. 964–968, 1989.

[124] M. Defrise, "A factorization method for the 3D X-ray transform," *Inverse Probl.*, vol. 11, no. 5, pp. 983–994, 1995.

[125] C. Comtat, P. Kinahan, M. Defrise *et al.*, "Fast Reconstruction of 3D PET Data with accurate statistical modeling," *IEEE Trans. Nucl. Sci.*, vol. 45, no. 3, pp. 1083–1089, 1998.

[126] A. J. Reader and J. Verhaeghe, "4D image reconstruction for emission tomography," *Phys. Med. Biol.*, vol. 59, no. 22, pp. R371–R418, 2014.

[127] J. Radon, "On the Determination of Functions From Their Integral Values Along Certain Manifolds *(translation)*," *IEEE Trans. Med. Imaging*, vol. 5, no. 4, pp. 170–176, 1986.

[128] R. L. Siddon, "Fast calculation of the exact radiological path for a three dimensional CT array," *Med. Phys.*, vol. 12, no. 2, pp. 252–255, 1985.

[129] S. Stute and C. Comtat, "Practical considerations for image-based PSF and blobs reconstruction in PET," *Phys. Med. Biol.*, vol. 58, no. 11, pp. 3849–3870, 2013.

[130] T. F. Budinger, "PET Instrumentation: What are the Limits?" *Semin. Nucl. Med.*, vol. 28, no. 3, pp. 247–267, 1998.

[131] G. L. Zeng, "Noise-weighted spatial domain FBP algorithm," *Med. Phys.*, vol. 41, no. 5, p. 051906, 2014.

[132] K. Van Slambrouck, S. Stute, C. Comtat *et al.*, "Bias Reduction for Low-Statistics PET: Maximum Likelihood Reconstruction With a Modified Poisson Distribution," *IEEE Trans. Med. Imaging*, vol. 34, no. 1, pp. 126–136, 2015.

[133] A. Boudjelal, Z. Messali, A. Elmoataz, and B. Attallah, "Improved Simultaneous Algebraic Reconstruction Technique Algorithm for Positron-Emission Tomography Image Reconstruction via Minimizing the Fast Total Variation," *J. Med. Imaging Radiat. Sci.*, vol. 48, no. 4, pp. 385–393, 2017.

[134] B. A. Ardekani, M. Braun, B. F. Hutton *et al.*, "Minimum cross-entropy reconstruction of PET images using prior anatomical information," *Phys. Med. Biol.*, vol. 41, no. 11, pp. 2497–2517, 1996.

[135] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Regularized Emission Image Reconstruction Using Imperfect Side Information," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 5, pp. 1464–1471, 1992.

[136] V. Y. Panin, G. L. Zeng, and G. T. Gullberg, "Total Variation Regulated EM Algorithm," *IEEE Trans. Nucl. Sci.*, vol. 46, no. 6, pp. 2202–2210, 1999.

[137] G. Wang and J. Qi, "Edge-Preserving PET Image Reconstruction Using Trust Optimization Transfer," *IEEE Trans. Med. Imaging*, vol. 34, no. 4, pp. 930–939, 2015.

[138] Z. Zhang, J. Ye, B. Chen *et al.*, "Investigation of optimization-based reconstruction with an image-total-variation constraint in PET," *Phys. Med. Biol.*, vol. 61, no. 16, pp. 6055–6084, 2016.

[139] Y. Teng, H. Sun, C. Guo, and Y. Kang, "ADMM-EM Method for L1-Norm Regularized Weighted Least Squares PET Reconstruction," *Comput. Math. Methods Med.*, vol. 2016, p. 6458289, 2016.

[140] X. Yu, C. Wang, H. Hu, and H. Liu, "Low Dose PET Image Reconstruction with Total Variation Using Alternating Direction Method," *PLoS One*, vol. 11, no. 12, p. 0166871, 2016.

[141] P. J. Huber, "Robust Estimation of a Location Parameter," *Ann. Math. Stat.*, vol. 35, no. 1, pp. 73–101, 1964.

[142] G. Wang and J. Qi, "Penalized Likelihood PET Image Reconstruction Using Patch-Based Edge-Preserving Regularization," *IEEE Trans. Med. Imaging*, vol. 31, no. 12, pp. 2194–2204, 2012.

[143] P. J. Green, "Bayesian Reconstructions From Emission Tomography Data Using a Modified EM Algorithm," *IEEE Trans. Med. Imaging*, vol. 9, no. 1, pp. 84–93, 1990.

[144] L. Lu, J. Ma, Q. Feng *et al.*, "Anatomy-guided brain PET imaging incorporating a joint prior model," *Phys. Med. Biol.*, vol. 60, no. 6, pp. 2145–2166, 2015.

[145] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, "Fast Gradient-Based Methods for Bayesian Reconstruction of Transmission and Emission PET Images," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 687–701, 1994.

[146] C. A. Bouman and K. Sauer, "A Unified Approach to Statistical Tomography Using Coordinate Descent Optimization," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 480–492, 1996.

[147] G. Schramm, M. Holler, A. Rezaei *et al.*, "Evaluation of Parallel Level Sets and Bowsher's Method as Segmentation-Free Anatomical Priors for Time-of-Flight PET Reconstruction," in *IEEE Trans. Med. Imaging*, vol. 37, no. 2, 2018, pp. 590–603.

[148] J. Rasch, E. M. Brinkmann, and M. Burger, "Joint reconstruction via coupled Bregman iterations with applications to PET-MR imaging," *Inverse Probl.*, vol. 34, no. 1, p. 014001, 2018.

[149] L. A. Shepp and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Trans. Med. Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

[150] K. Lange and R. Carson, "EM Reconstruction Algorithms for Emission and Transmission Tomography," *J. Comput. Assist. Tomogr.*, vol. 8, no. 2, pp. 306–316, 1984.

[151] P. J. Green, "On Use of the EM for Penalized Likelihood Estimation," *J. R. Stat. Soc. B*, vol. 52, no. 3, pp. 443–452, 1990.

[152] K. Thielemans, C. Tsoumpas, S. Mustafovic *et al.*, "STIR: software for tomographic image reconstruction release 2," *Phys. Med. Biol.*, vol. 57, no. 4, pp. 867–883, 2012.

[153] A. R. De Pierro, "A Modified Expectation Maximization Algorithm for Penalized Likelihood Estimation in Emission Tomography," *IEEE Trans. Med. Imaging*, vol. 14, no. 1, pp. 132–137, 1995.

[154] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, 2006.

[155] B. Bai, Q. Li, and R. M. Leahy, "MR-Guided PET Image Reconstruction," *Semin. Nucl. Med.*, vol. 43, no. 1, pp. 30–44, 2013.

[156] K. Erlandsson, I. Buvat, P. H. Pretorius *et al.*, "A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology," *Phys. Med. Biol.*, vol. 57, no. 21, pp. R119–R159, 2012.

[157] J. Yan, J. C.-S. Lim, and D. W. Townsend, "MRI-guided brain PET image filtering and partial volume correction," *Phys. Med. Biol.*, vol. 60, no. 3, pp. 961–976, 2015.

[158] G. Gindi, M. Lee, A. Rangarajan, and I. G. Zubal, "Bayesian Reconstruction of Functional Images Using Anatomical Information as Priors," *IEEE Trans. Med. Imaging*, vol. 12, no. 4, pp. 670–680, 1993.

[159] X. Ouyang, W. H. Wong, V. E. Johnson *et al.*, "Incorporation of Correlated Structural Images in PET Image Reconstruction," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 627–640, 1994.

[160] B. Lipinski, H. Herzog, E. Rota Kops *et al.*, "Expectation Maximization Reconstruction of Positron Emission Tomography Images Using Anatomical Magnetic Resonance Information," *IEEE Trans. Med. Imaging*, vol. 16, no. 2, pp. 129–136, 1997.

[161] A. O. Hero, R. Piramuthu, J. A. Fessier, and S. R. Titus, "Minimax Emission Computed Tomography Using High-Resolution Anatomical Side Information and B-Spline Models," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 920–938, 1999.

[162] C. Comtat, P. E. Kinahan, J. A. Fessler *et al.*, "Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels," *Phys. Med. Biol.*, vol. 47, no. 1, pp. 1–20, 2002.

[163] A. M. Alessio and P. E. Kinahan, "Improved quantitation for PET/CT image reconstruction with system modeling and anatomical priors," *Med. Phys.*, vol. 33, no. 11, pp. 4095–4103, 2006.

[164] J. Cheng-Liao and J. Qi, "PET image reconstruction with anatomical edge guided level set prior," *Phys. Med. Biol.*, vol. 56, no. 21, pp. 6899–6918, 2011.

[165] J. E. Bowsher, V. E. Johnson, T. G. Turkington *et al.*, "Bayesian Reconstruction and Use of Anatomical *A Priori* Information for Emission Tomography," *IEEE Trans. Med. Imaging*, vol. 15, no. 5, pp. 673–686, 1996.

[166] S. Sastry and R. E. Carson, "Multimodality Bayesian Algorithm for Image Reconstruction in Positron Emission Tomography: A Tissue Composition Model," *IEEE Trans. Med. Imaging*, vol. 16, no. 6, pp. 750–61, 1997.

[167] K. Baete, J. Nuyts, W. Van Paesschen *et al.*, "Anatomical-Based FDG-PET Reconstruction for the Detection of Hypo-Metabolic Regions in Epilepsy," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 510–519, 2004.

[168] K. Vunckx, A. Atre, K. Baete *et al.*, "Evaluation of Three MRI-Based Anatomical Priors for Quantitative PET Brain Imaging," *IEEE Trans. Med. Imaging*, vol. 31, no. 3, pp. 599–612, 2012.

[169] J. E. Bowsher, H. Yuan, L. W. Hedlund *et al.*, "Utilizing MRI Information to Estimate F18-FDG Distributions in Rat Flank Tumors," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 4, 2004, pp. 2488–2492.

[170] V. G. Nguyen and S.-J. Lee, "Incorporating Anatomical Side Information into PET Reconstruction Using Nonlocal Regularization," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3961–3973, 2013.

[171] R. Loeb, N. Navab, and S. I. Ziegler, "Direct Parametric Reconstruction Using Anatomical Regularization for Simultaneous PET/MRI Data," *IEEE Trans. Med. Imaging*, vol. 34, no. 11, pp. 2233–2247, 2015.

[172] A. Mehranian, M. A. Belzunce, F. Niccolini *et al.*, "PET image reconstruction using multi-parametric anato-functional priors," *Phys. Med. Biol.*, vol. 62, no. 15, pp. 5975–6007, 2017.

[173] M. J. Ehrhardt, P. Markiewicz, M. Liljeroth *et al.*, "PET Reconstruction with an Anatomical MRI Prior Using Parallel Level Sets," *IEEE Trans. Med. Imaging*, vol. 35, no. 9, pp. 2189–2199, 2016.

[174] A. Rangarajan, I.-T. Hsiao, and G. Gindi, "A Bayesian Joint Mixture Framework for the Integration of Anatomical Information in Functional Image Reconstruction," *J. Math. Imaging Vis.*, vol. 12, no. 3, pp. 199–217, 2000.

[175] S. Somayajula, E. Asma, and R. Leahy, "PET Image Reconstruction Using Anatomical Information Through Mutual Information Based Priors," in *IEEE Nucl. Sci. Symp. Conf. Rec.* IEEE, 2005, pp. 2722–2726.

[176] J. Nuyts, "The Use of Mutual Information and Joint Entropy for Anatomical Priors in Emission Tomography," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 6. IEEE, 2007, pp. 4149–4154.

[177] J. Tang and A. Rahmim, "Bayesian PET image reconstruction incorporating anato-functional joint entropy," *Phys. Med. Biol.*, vol. 54, no. 23, pp. 7063–7075, 2009.

[178] S. Somayajula, C. Panagiotou, A. Rangarajan *et al.*, "PET Image Reconstruction Using Information Theoretic Anatomical Priors," *IEEE Trans. Med. Imaging*, vol. 30, no. 3, pp. 537–549, 2011.

[179] J. Tang and A. Rahmim, "Anatomy assisted PET image reconstruction incorporating multi-resolution joint entropy," *Phys. Med. Biol.*, vol. 60, no. 1, pp. 31–48, 2015.

[180] G. Wang and J. Qi, "PET Image Reconstruction Using Kernel Method," *IEEE Trans. Med. Imaging*, vol. 34, no. 1, pp. 61–71, 2015.

[181] W. Hutchcroft, G. Wang, K. T. Chen *et al.*, "Anatomically-aided PET reconstruction using the kernel method," *Phys. Med. Biol.*, vol. 61, no. 18, pp. 6668–6683, 2016.

[182] P. Novosad and A. J. Reader, "MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions," *Phys. Med. Biol.*, vol. 61, no. 12, pp. 4624–4644, 2016.

[183] K. Gong, J. Cheng-Liao, G. Wang *et al.*, "Direct Patlak Reconstruction From Dynamic PET Data Using the Kernel Method with MRI Information Based on Structural Similarity," *IEEE Trans. Med. Imaging*, vol. 37, no. 4, pp. 955–964, 2018.

[184] D. Deidda, N. Karakatsanis, P. M. Robson *et al.*, "Hybrid PET-MR List-Mode Kernelized Expectation Maximization Reconstruction for Quantitative PET Images of the Carotid Arteries," *IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2017.

[185] J. Bland, A. Mehranian, M. A. Belzunce *et al.*, "MR-Guided Kernel EM Reconstruction for Reduced Dose PET Imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 3, pp. 235–243, 2018.

[186] J. Bland, M. A. Belzunce, S. Ellis *et al.*, "Spatially-Compact MR-Guided Kernel EM for PET Image Reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 5, pp. 470–482, 2018.

[187] S. Chen, H. Liu, P. Shi, and Y. Chen, "Sparse representation and dictionary learning penalized image reconstruction for positron emission tomography," *Phys. Med. Biol.*, vol. 60, no. 2, pp. 807–823, 2015.

[188] M. S. Tahaei and A. J. Reader, "Patch-based image reconstruction for PET using prior-image derived dictionaries," *Phys. Med. Biol.*, vol. 61, no. 18, pp. 6833–6855, 2016.

[189] J. Nuyts, K. Baete, D. Bequé, and P. Dupont, "Comparison Between MAP and Postprocessed ML for Image Reconstruction in Emission Tomography When Anatomical Knowledge Is Available," *IEEE Trans. Med. Imaging*, vol. 24, no. 5, pp. 667–675, 2005.

[190] A. Buades, B. Coll, and J. M. Morel, "A Review of Image Denoising Algorithms, with a New One," *SIAM J. Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[191] C. Song, Y. Yang, P. H. Pretorius, and M. A. King, "4D non-local means post-filtering for cardiac gated SPECT," *Phys. Med. Biol.*, vol. 63, no. 3, p. 035026, 2018.

[192] C. Chan, R. Fulton, R. Barnett *et al.*, "Postreconstruction Nonlocal Means Filtering of Whole-Body PET with an Anatomical Prior," *IEEE Trans. Med. Imaging*, vol. 33, no. 3, pp. 636–650, 2014.

[193] S. Y. Chun, J. A. Fessler, and Y. K. Dewaraja, "Post-reconstruction non-local means filtering methods using CT side information for quantitative SPECT," *Phys. Med. Biol.*, vol. 58, no. 17, pp. 6225–6240, 2013.

[194] G.-H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med. Phys.*, vol. 35, no. 2, pp. 660–663, 2008.

[195] D. Donoho, "Compressed Sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[196] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The Application of Compressed Sensing for Rapid MR Imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[197] P. T. Lauzier, J. Tang, and G. H. Chen, "Prior image constrained compressed sensing: Implementation and performance evaluation," *Med. Phys.*, vol. 39, no. 1, pp. 66–80, 2012.

[198] J. C. Ramirez-Giraldo, J. Trzasko, S. Leng *et al.*, "Nonconvex prior image constrained compressed sensing (NCPICCS): Theory and simulations on perfusion CT," *Med. Phys.*, vol. 38, no. 4, pp. 2157–2167, 2011.

[199] J. W. Stayman, H. Dang, Y. Ding, and J. H. Siewerdsen, "PIRPLE: A penalized-likelihood framework for incorporation of prior images in CT reconstruction," *Phys. Med. Biol.*, vol. 58, no. 21, pp. 7563–7582, 2013.

[200] H. Dang, A. S. Wang, M. S. Sussman *et al.*, "dPIRPLE: a joint estimation framework for deformable registration and penalized-likelihood CT image reconstruction using prior images," *Phys. Med. Biol.*, vol. 59, no. 17, pp. 4799–4826, 2014.

[201] A. Pourmorteza, H. Dang, J. H. Siewerdsen, and J. W. Stayman, "Reconstruction of difference in sequential CT studies using penalized likelihood estimation," *Phys. Med. Biol.*, vol. 61, no. 5, pp. 1986–2002, 2016.

[202] S. Seyyedi, E. Liapi, T. Lasser *et al.*, "Low-Dose CT Perfusion of the Liver using Reconstruction of Difference," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 3, pp. 205–214, 2018.

[203] P. Wu, J. W. Stayman, M. Mow *et al.*, "Reconstruction-of-difference (RoD) imaging for cone-beam CT neuro-angiography," *Phys. Med. Biol.*, vol. 63, no. 11, p. 115004, 2018.

[204] M. Lustig, J. M. Santos, D. Donoho, and J. M. Pauly, "k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity," in *Proc. 14th Annu. Meet. Int. Soc. Magn. Reson. Med.*, 2006, p. 2420.

[205] B. Tremoulhéac, N. Dikaios, D. Atkinson, and S. R. Arridge, "Dynamic MR Image Reconstruction-Separation From Undersampled (k,t)-Space via Low-Rank Plus Sparse Prior," *IEEE Trans. Med. Imaging*, vol. 33, no. 8, pp. 1689–1701, 2014.

[206] S. F. Roohi, D. Zonoobi, A. A. Kassim, and J. L. Jaremko, "Multi-dimensional low rank plus sparse decomposition for reconstruction of under-sampled dynamic MRI," *Pattern Recognit.*, vol. 63, pp. 667–679, 2017.

[207] L. Feng, M. B. Srichai, R. P. Lim *et al.*, "Highly Accelerated Real-Time Cardiac Cine MRI Using k–t SPARSE-SENSE," *Magn. Reson. Med.*, vol. 70, no. 1, pp. 64–74, 2013.

[208] L. Feng, L. Axel, H. Chandarana *et al.*, "XD-GRASP: Golden-Angle Radial MRI with Reconstruction of Extra Motion-State Dimensions Using Compressed Sensing," *Magn. Reson. Med.*, vol. 75, no. 2, pp. 775–788, 2016.

[209] L. Feng, S. Coppo, D. Piccini *et al.*, "5D Whole-Heart Sparse MRI," *Magn. Reson. Med.*, vol. 79, no. 2, pp. 826–838, 2018.

[210] T. Correia, G. Ginami, G. Cruz *et al.*, "Optimized respiratory-resolved motion-compensated 3D Cartesian coronary MR angiography," *Magn. Reson. Med.*, 2018.

[211] L. Weizman, Y. C. Eldar, and D. Ben Bashat, "Compressed sensing for longitudinal MRI: An adaptive-weighted approach," *Med. Phys.*, vol. 42, no. 9, pp. 5195–5208, 2015.

[212] R. Kang, G. Qu, B. Cao, and L. Yan, "Combined Similarity to Reference Image with Joint Sparsifying Transform for Longitudinal Compressive Sensing MRI," *Math. Probl. Eng.*, vol. 2016, p. 4162194, 2016.

[213] Y. Rakvongthai, F. Fahey, K. Borvorntanajanya *et al.*, "Joint reconstruction of Ictal/inter-ictal SPECT data for improved epileptic foci localization," *Med. Phys.*, vol. 44, no. 4, pp. 1437–1444, 2017.

[214] A. Ihsani, J. Dutta, K. Johnson *et al.*, "A Differential PET Image Reconstruction Method for Improved Sensitivity of Tau Protein Deposition in Alzheimer Disease Progression Monitoring," in *IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2017.

[215] X. He, L. Cheng, J. A. Fessler, and E. C. Frey, "Regularized Image Reconstruction Algorithms for Dual-Isotope Myocardial Perfusion SPECT (MPS) Imaging Using a Cross-Tracer Prior," *IEEE Trans. Med. Imaging*, vol. 30, no. 6, pp. 1169–1183, 2011.

[216] H.-P. W. Schlemmer, B. J. Pichler, M. Schmand *et al.*, "Simultaneous MR/PET Imaging of the Human Brain: Feasibility Study," *Radiology*, vol. 248, no. 3, pp. 1028–1035, 2008.

[217] G. Delso, S. Fürst, B. Jakoby *et al.*, "Performance Measurements of the Siemens mMR Integrated Whole-Body PET/MR Scanner," *J. Nucl. Med.*, vol. 52, no. 12, pp. 1914–1922, 2011.

[218] J. Tsao and S. Kozerke, "MRI temporal acceleration techniques," *J. Magn. Reson. Imaging*, vol. 36, no. 3, pp. 543–560, 2012.

[219] K. G. Hollingsworth, "Reducing acquisition time in clinical MRI by data undersampling and compressed sensing reconstruction," *Phys. Med. Biol.*, vol. 60, no. 21, pp. R297–R322, 2015.

[220] M. J. Ehrhardt, K. Thielemans, L. Pizarro *et al.*, "Joint reconstruction of PET-MRI by exploiting structural similarity," *Inverse Probl.*, vol. 31, no. 1, p. 15001, 2015.

[221] F. Knoll, M. Holler, T. Koesters *et al.*, "Joint MR-PET Reconstruction Using a Multi-Channel Image Regularizer," *IEEE Trans. Med. Imaging*, vol. 36, no. 1, pp. 1–16, 2017.

[222] A. Mehranian, M. A. Belzunce, C. J. Mcginnity, and C. Prieto, "Multimodal weighted quadratic priors for robust intensity independent synergistic PET-MR reconstruction," in *IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2017, pp. 1–25.

[223] A. Mehranian, M. A. Belzunce, C. Prieto *et al.*, "Synergistic PET and SENSE MR Image Reconstruction Using Joint Sparsity Regularization," *IEEE Trans. Med. Imaging*, vol. 37, no. 1, pp. 20–34, 2018.

[224] Y. Zhang and X. Zhang, "PET-MRI joint reconstruction with common edge weighted total variation regularization," *Inverse Probl.*, vol. 34, no. 6, p. 065006, 2018.

[225] S. Ellis and A. J. Reader, "Simultaneous maximum a posteriori longitudinal PET image reconstruction," *Phys. Med. Biol.*, vol. 62, no. 17, pp. 6963–6979, 2017.

[226] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization," *Proc. Natl. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.

[227] A. Rahmim, K. Dinelle, J. C. Cheng *et al.*, "Accurate Event-Driven Motion Compensation in High-Resolution PET Incorporating Scattered and Random Events," *IEEE Trans. Med. Imaging*, vol. 27, no. 8, pp. 1018–1033, 2008.

[228] M. A. Belzunce, J. O'Doherty, and A. J. Reader, "Impact of Axial Compression for the mMR Simultaneous PET-MR Scanner," in *2015 IEEE Nucl. Sci. Symp. Med. Imaging Conf.*, 2015, pp. 1–3.

[229] M. A. Belzunce and A. J. Reader, "Assessment of the impact of modeling axial compression on PET image reconstruction," *Med. Phys.*, vol. 44, no. 10, pp. 5172–5186, 2017.

[230] M. W. Jacobson and J. A. Fessler, "Joint Estimation of Image and Deformation Parameters in Motion-Corrected PET," in *2003 IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 5, 2003, pp. 3290–3294.

[231] M. Blume, A. Martinez-Möller, A. Keil *et al.*, "Joint Reconstruction of Image and Motion in Gated Positron Emission Tomography," *IEEE Trans. Med. Imaging*, vol. 29, no. 11, pp. 1892–1906, 2010.

[232] M. Blume, N. Navab, and M. Rafecas, "Joint image and motion reconstruction for PET using a B-spline motion model," *Phys. Med. Biol.*, vol. 57, no. 24, pp. 8249–8270, 2012.

[233] F. Kalantari, T. Li, M. Jin, and J. Wang, "Respiratory motion correction in 4D-PET by simultaneous motion estimation and image reconstruction (SMEIR)," *Phys. Med. Biol.*, vol. 61, no. 15, pp. 5639–5661, 2016.

[234] S. Ellis and A. J. Reader, "Penalized maximum likelihood simultaneous longitudinal PET image reconstruction with difference-image priors," *Med. Phys.*, vol. 45, no. 7, pp. 3001–3018, 2018.

[235] L. Wasserman, *All of Non Parametric Statistics*. New York: Springer, 2005.

[236] A. Mehranian, F. Kotasidis, and H. Zaidi, "Accelerated time-of-flight (TOF) PET image reconstruction using TOF bin subsetization and TOF weighting matrix pre-computation," *Phys. Med. Biol.*, vol. 61, no. 3, p. 1309, 2016.

[237] H. M. Hudson and R. S. Larkin, "Accelerated Image Reconstruction Using Ordered Subsets of Projection Data," *IEEE Trans. Med. Imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[238] B. W. Jakoby, Y. Bercier, M. Conti *et al.*, "Physical and clinical performance of the mCT time-of-flight PET/CT scanner," *Phys. Med. Biol.*, vol. 56, no. 8, pp. 2375–2389, 2011.

[239] F. P. Oliveira and J. M. R. Tavares, "Medical image registration: A review," *Comput. Methods Biomech. Biomed. Engin.*, vol. 17, no. 2, pp. 73–93, 2014.

[240] S. Ellis, A. Mallia, C. J. McGinnity *et al.*, "Multi-Tracer Guided PET Image Reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 5, pp. 499–509, 2018.

[241] A. Marusyk and K. Polyak, "Tumor heterogeneity: causes and consequences." *Biochim. Biophys. Acta*, vol. 1805, no. 1, pp. 105–17, 2010.

[242] I. G. Zubal, S. S. Spencer, K. Imam *et al.*, "Difference Images Calculated from Ictal and Interictal Technetium-99m-HMPAO SPECT Scans of Epilepsy," *J. Nucl. Med.*, vol. 36, no. 4, pp. 684–689, 1995.

[243] A. Desai, K. Bekelis, V. M. Thadani *et al.*, "Interictal PET and ictal subtraction SPECT: Sensitivity in the detection of seizure foci in patients with medically intractable epilepsy," *Epilepsia*, vol. 54, no. 2, pp. 341–350, 2013.

[244] A. Perissinotti, X. Setoain, J. Aparicio *et al.*, "Clinical Role of Subtraction Ictal SPECT Coregistered to MR Imaging and 18F-FDG PET in Pediatric Epilepsy," *J. Nucl. Med.*, vol. 55, no. 7, pp. 1099–1105, 2014.

[245] T.-S. Lee, W. P. Segars, and B. M. W. Tsui, "Study of Parameters Characterizing Space-Time Gibbs Priors for 4D MAP-RBI-EM in Gated Myocardial Perfusion SPECT," in *IEEE Nucl. Sci. Symp. Conf. Rec.*, vol. 4. IEEE, 2005, pp. 2124–2128.

[246] C. Tsoumpas, C. Buerger, P. Mollet, and P. K. Marsden, "Fast analytic simulation toolkit for generation of 4D PET-MR data from real dynamic MR acquisitions," *J. Phys.: Conf. Ser.*, vol. 317, p. 012020, 2011.

Appendix: Derivation of Gradient Terms for Difference-Image Priors

Given the set of longitudinal images $\{\boldsymbol{\theta}_s\}$, the general difference-image prior is defined as a summation of contributions from each difference image, i.e.

$$U(\{\boldsymbol{\theta}_s\}) = \sum_{s=1}^{S}\sum_{k=1}^{S} w_{sk} u(\boldsymbol{\delta}_{sk}). \tag{A.1}$$

where $u(\cdot)$ is some function applied to the difference images, and the difference images are corrected for varying activity levels and misalignment according to:

$$\boldsymbol{\delta}_{sk} = \alpha_k \boldsymbol{M}_k \boldsymbol{\theta}_k - \alpha_s \boldsymbol{M}_s \boldsymbol{\theta}_s, \tag{A.2}$$

where α_s is a scaling factor to account for the activity level in scan s and $\boldsymbol{M}_s$ is a spatial operator that transforms the longitudinal image $\boldsymbol{\theta}_s$ into the single space where all difference images are calculated.

Therefore, the derivative with respect to the jth voxel value in $\boldsymbol{\theta}_b$ can be written as:

$$\begin{aligned}\frac{\partial U}{\partial \theta_{j,b}} &= \frac{\partial}{\partial \theta_{j,b}} \sum_{s=1}^{S}\sum_{k=1}^{S} w_{sk} u(\boldsymbol{\delta}_{sk}) \\ &= \sum_{s=1}^{S} w_{sb} \frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \theta_{j,b}} + \sum_{k=1}^{S} w_{bk} \frac{\partial u(\boldsymbol{\delta}_{bk})}{\partial \theta_{j,b}}.\end{aligned} \tag{A.3}$$

Here the total derivative has been split into two contributing terms where $k = b$ and where $s = b$ (since $s = k = b$ has no contribution to the gradient

because the difference image is always zero in these cases).

The chain rule can then be used on each of the two terms in Equation A.3, for example:

$$\frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \theta_{j,b}} = \sum_{l=1}^{J} \frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \delta_{l,sb}} \frac{\partial \delta_{l,sb}}{\partial \theta_{j,b}}. \tag{A.4}$$

The chain rule can be used again, letting $\boldsymbol{v}_b = \boldsymbol{M}_b \boldsymbol{\theta}_b$, to get

$$\frac{\partial \delta_{l,sb}}{\partial \theta_{j,b}} = \frac{\partial \delta_{l,sb}}{\partial v_{l,b}} \frac{\partial v_{l,b}}{\partial \theta_{j,b}}. \tag{A.5}$$

Now both terms in Equation A.5 can be calculated. Firstly,

$$\frac{\partial \delta_{l,sb}}{\partial v_{l,b}} = \alpha_b. \tag{A.6}$$

Secondly, since

$$v_{l,b} = \sum_{j=1}^{J} M_{lj,b} \theta_{j,b} \tag{A.7}$$

then

$$\frac{\partial v_{l,b}}{\partial \theta_{j,b}} = M_{lj,b}. \tag{A.8}$$

Therefore, bringing it all together,

$$\begin{aligned} \frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \theta_{j,b}} &= \sum_{l=1}^{J} \frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \delta_{l,sb}} \cdot \alpha_b \cdot M_{lj,b} \\ &= \alpha_b \sum_{l=1}^{J} M_{lj,b} \frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \delta_{l,sb}}. \end{aligned} \tag{A.9}$$

Given the gradient vector,

$$\nabla u_{ks} = \left[\frac{\partial u(\boldsymbol{\delta}_{ks})}{\partial \delta_{1,ks}}, ..., \frac{\partial u(\boldsymbol{\delta}_{ks})}{\partial \delta_{J,ks}} \right]^T, \tag{A.10}$$

Equation A.9 can be written as:

$$\frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \theta_{j,b}} = \alpha_b \left[\boldsymbol{M}_b^T \nabla u_{sb} \right]_j. \tag{A.11}$$

where $[\cdot]_j$ denotes the jth element of a vector expression.

Using an identical argument, the remaining term in Equation A.3 can be shown to be given by

$$\frac{\partial u(\boldsymbol{\delta}_{bk})}{\partial \theta_{j,b}} = -\alpha_b \left[\boldsymbol{M}_b^T \nabla u_{bk}\right]_j, \tag{A.12}$$

where the negative arises from the reversing of the order of the images in the definition of the difference image.

Therefore, the total gradient of the penalty can be written as:

$$\begin{aligned}\frac{\partial U}{\partial \theta_{j,b}} &= \sum_{s=1}^{S} w_{sb} \frac{\partial u(\boldsymbol{\delta}_{sb})}{\partial \theta_{j,b}} + \sum_{k=1}^{S} w_{bk} \frac{\partial u(\boldsymbol{\delta}_{bk})}{\partial \theta_{j,b}} \\ &= \sum_{s=1}^{S} w_{sb}\alpha_b \left[\boldsymbol{M}_b^T \nabla u_{sb}\right]_j - \sum_{k=1}^{S} w_{bk}\alpha_b \left[\boldsymbol{M}_b^T \nabla u_{bk}\right]_j\end{aligned} \tag{A.13}$$

Now, since $\boldsymbol{\delta}_{bk} = -\boldsymbol{\delta}_{kb}$ and the derivative is a linear operator, it is apparent that $\nabla u_{bk} = -\nabla u_{kb}$. Furthermore, since both indices s and k sum over all contributions from all longitudinal scans, k can be replaced by s.

Therefore, the total derivative is given by

$$\begin{aligned}\frac{\partial U}{\partial \theta_{j,b}} &= \sum_{s=1}^{S} w_{sb}\alpha_b \left[\boldsymbol{M}_b^T \nabla u_{sb}\right]_j + \sum_{s=1}^{S} w_{bs}\alpha_b \left[\boldsymbol{M}_b^T \nabla u_{sb}\right]_j \\ &= \sum_{s=1}^{S} \alpha_b (w_{sb} + w_{bs}) \left[\boldsymbol{M}_b^T \nabla u_{sb}\right]_j\end{aligned} \tag{A.14}$$

Finally, since w_{sb} and w_{bs} represent the coupling between scans s and b, it is reasonable to constrain them to be equal. This gives the penalty gradient as:

$$\frac{\partial U}{\partial \theta_{j,b}} = 2\alpha_b \sum_{s=1}^{S} w_{sb} \left[\boldsymbol{M}_b^T \nabla u_{sb}\right]_j \tag{A.15}$$

This is the gradient of a general difference-image prior given in Equation 4.4, as used in Chapter 4.